



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Measures to Enhance the Effectiveness of International Climate Agreements

**Citation for published version:**

Al Khourdajie, A & Finus, M 2018 'Measures to Enhance the Effectiveness of International Climate Agreements: The Case of Border Carbon Adjustments' Bath Economics Research Papers.  
<[http://www.bath.ac.uk/economics/research/working-papers/2018-papers/Working\\_Paper\\_71-18\\_MichaelFinus.pdf](http://www.bath.ac.uk/economics/research/working-papers/2018-papers/Working_Paper_71-18_MichaelFinus.pdf)>

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Other version

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



Measures to Enhance the Effectiveness of International Climate  
Agreements: The Case of Border Carbon Adjustments

Alaa Al Khourdajie and Michael Finus  
No. 71/18

**BATH ECONOMICS RESEARCH PAPERS**

**Department of Economics**

Department of  
Economics



UNIVERSITY OF  
**BATH**

# Measures to Enhance the Effectiveness of International Climate Agreements: The Case of Border Carbon Adjustments

Alaa Al Khourdajie\* and Michael Finus†

## Abstract

Unilateral or sub-global actions on climate change are not very effective but global action is not achievable due to strong free-rider incentives. These incentives arise because non-signatories benefit from emission reductions of signatories without incurring abatement costs. Moreover, non-signatories may even increase their emissions as a result of lower fuel prices and because of a relocation of production from “clean” signatory to “dirty” non-signatory countries, typically referred to as carbon leakage. We study whether and under which conditions border carbon adjustments (BCAs) can mitigate free-riding in a simple strategic trade model which captures consumers’ taste for variety. We show that BCAs lead to larger stable international climate agreements associated with global welfare gains but only if treaties are of the open membership type and do not serve the interests of few countries which may prefer an exclusive membership rule. We focus on the strategic interaction between signatory and non-signatory countries and the game-theoretic properties and institutional rules of treaty formation.

**Keywords:** self-enforcing international climate agreements, international trade, border carbon adjustments, consumers’ taste for variety and horizontal product differentiation.

**JEL-Classification:** C71, D62, F18, H23, H41, Q54.

---

\*Email: Alaa.Alkhourdajie@ed.ac.uk | School of Economics, University of Edinburgh, Edinburgh, EH8 9JT, UK.

†Corresponding author: Email: M.Finus@bath.ac.uk | Department of Economics, University of Bath, Claverton Down, Bath, BA2 7AY, UK.

# 1 Introduction

Unilateral or sub-global actions on climate change are not very effective but global action is not achievable due to strong free-rider incentives. These incentives arise because non-signatories benefit from emission reductions of signatories without incurring abatement costs. Moreover, non-signatories may even increase their emissions due to higher fuel consumption as a result of lower fuel prices and because of a relocation of production from “clean” signatory to “dirty” non-signatory countries, typically referred to as carbon leakage. In order to mitigate these free-rider incentives and to make climate agreements more effective, trade measures, like border carbon adjustments (BCAs), have become increasingly popular in the policy debate (e.g. Financial Times, Feb. 12, 2017 and Zenghelis and Stern 2009) but also in academic circles (e.g. Böhringer et al. 2017b, Fischer and Fox 2012 and Keen and Kotsogiannis 2014). Border carbon adjustments<sup>1</sup> is a complementary measure by a group of countries, which we call signatory countries or members to a climate treaty, which pursue a more ambitious climate policy than other countries, which we call non-signatory countries or non-members, whereby a tariff, which is de facto an emission tax, is imposed on imports. At least in theory, the tariff is a tax per unit of emissions, which is the same as the price of carbon (resulting from any form of environmental regulation) in signatory countries. Most authors talk about full BCAs if tariffs are accompanied by output-based rebates (OBRs) whereby firms producing under stricter environmental standards receive rebates for their exports. OBRs can be paid directly or indirectly like the exemption of emission-intensive and trade exposed industries from strict environmental regulation, like the steel and aluminum industry within the EU-ETS, e.g. through a generous free allocation of tradeable permits (Böhringer et al. 2017b). An OBR on its own is typically viewed as an anti-leakage measure, but it violates the strong polluters pay principle and, at least ceteris paribus, is less effective in reducing emissions than tariffs. With OBRs, BCAs are de facto a consumption-based environmental tax imposed in signatory markets.

The academic discussion about BCAs has broadly focused on three issues. The first issue is about the detailed design and practicability of trade measures for environmental reasons and to which extent they are compatible with WTO-rules (e.g. Fischer 2015, Fischer and Fox 2012 and Mehling et al. 2017). It appears that the overarching conclusion is that, despite many practical obstacles of implementation, BCAs do not generally violate WTO-rules and are a good second-best instrument in the absence of a global climate treaty.

The second issue is about the economic justification of trade measures for environmental reasons, which touches on the efficiency and costs of these measures.<sup>2</sup> For

---

<sup>1</sup>We use the term BCAs to stress the environmental justification of border tax adjustments (BTAs), which are also discussed for other reasons, like different labor standards and cooperate taxes across countries.

<sup>2</sup>For an excellent discussion of this issue, see for instance Helm et al. (2012).

instance, Stiglitz (2006) argues that the absence of carbon prices de facto constitute a subsidy to dirty production and hence to ensure a level playing field while internalizing the social cost of carbon, BCAs are just correcting a market imperfection. Despite conventional trade theory argues against trade-barriers because different environmental standards simply reflect different environmental preferences and/or comparative advantages related to the relative abundance of the environment and natural resources of countries, the correction of distortions through market interventions can be justified in the context of global pollutants. The internalization through first- and second-best instruments has been analyzed for instance by Copeland and Taylor (1995), Markusen (1975), Hoel (1996) and, more recently, e.g. by Keen and Kotsogiannis (2014), Vlassis (2014) and Tsakiris (2014). It is clear that if BCAs do not enforce uniformly higher environmental standards at the global scale but only among a subgroup of countries, the welfare gains from a tougher environmental policy may be small if not negative as BCAs may seriously harm outsiders. In other words, in an ideal word, BCAs are a threat to enforce full cooperation (or something close to this), but, if successful, are not implemented.

The third issue relates to the effectiveness of trade measures. The bulk of the literature conducts simulations with empirically calibrated computable general equilibrium (CGE) models in order to estimate the ex-ante leakage effect reduction of BCAs (full or partial) as well as OBRs, at the aggregate and in particular sectors (Böhringer et al. 2012, 2014, 2015, 2017a,b, Coron 2012 and Fischer and Fox 2012; see also Branger and Quirion 2014 for a meta-analysis of 25 studies). Most of these studies confirm the conclusion that OBRs can reduce carbon-leakage but only import tariffs lead also to a noticeable overall reduction of emissions as they curtail also consumption. Due to the complexity of these multi-country multi-sector models, typically, the environmental target/tax of signatory countries is set exogenously and the environmental standards of non-signatories is fixed at the level without BCAs. That is, the strategic interaction between signatories and non-signatories in a game-theoretic sense is not explicitly captured. Moreover, in contrast to the original idea of BCAs, namely to support a climate treaty, the formation and stability of climate treaties is not considered. An exception is Irfanoglu et al. (2015), but the analysis is restricted to three countries. Although conceptually interesting, Helm et al. (2012) remain at a rather stylized level in their analysis. Another exception is Weitzel et al. (2012) who analyze stability of coalitions with the DART-model with 10 world regions and 9 different sectors. However, their stability is seriously simplified as they consider only stability of the grand coalition and a group of annex 1 countries.

A smaller part of the literature on the third issue is theoretical in nature and has focused on the endogenous choice of policy levels among signatory countries, including the reaction by non-signatory countries, in a strategic trade model of imperfect competition. More surprisingly, also this literature has mainly ignored the issue of treaty formation by restricting attention to two countries (e.g. Baksi and Chaudhuri 2017, Eyland and Zaccour 2013 and 2014). Typically, these models consider a home

country, which recognizes or suffers more from environmental damages than a foreign country, with lower or zero environmental damages. Hence, initially, subsidies are lower or taxes higher in the home country in a non-cooperative equilibrium, typically viewed as the status quo. Full cooperation fails in the absence of transfer payments as the foreign countries welfare is lower under full cooperation than in the status quo if the difference in the evaluation of environmental damages is pronounced enough. It is then shown that BCAs may be able to lower the welfare of the foreign country sufficiently enough compared to the status quo such that the foreign country agrees to full cooperation. That is, BCAs act as a threat to enforce cooperation. In these simple models, an agreement is stable if it is individually rational compared to some status quo, and the BCA changes this status quo. A slight variation of this stability concept is considered in Anouliés (2014) who shows that the minimum discount factor necessary to enforce a fully cooperative solution in a repeated game is reduced if this triggers the implementation of the BCA-regime as punishment instead of the non-cooperative fall back position. Probably more interesting is the finding that for most parameter values the home country prefers the BCA-regime over the fully cooperative solution, suggesting that BCAs are a credible threat. As far as we are aware, only Baksi and Chaudhuri (2014) consider BCAs in n-country model of coalition formation, but they only analyze the stability of the grand coalition and ignore the entire process of treaty formation.

Our paper is in the tradition of the game-theoretic literature on the third issue. In contrast to previous papers, we model the entire process of treaty formation including stability, though admittedly, the model abstracts from many details captured by CGE models. We are interested in two main issues.

Firstly, how BCAs change the properties of an n-country public good provision game with the possibility of forming self-enforcing treaties as known from the literature on international environmental agreements.<sup>3</sup> We distinguish between positive and normative properties. Positive properties help to explain why without BCAs climate agreements are not very effective and the normative properties evaluate outcomes in relation to the benchmarks of a non-cooperative and fully cooperative outcome. Both type of properties give rise to what Barrett (1994a) called the paradox of cooperation: stable agreements are either small, shallow or both whenever the gains from cooperation would be large; they may enjoy large participation but only if gains from cooperation are small.<sup>4</sup> We show that BCAs change some of the standard properties fundamentally. On the one hand, countries find it less attractive to stay outside an agreement as they may now be negatively affected by other countries joining a climate agreement. Moreover, countries find it more attractive to join an agreement as the gains from cooperation are increased. On the other hand, a sequential enlargement of an agreement may no longer be associated with an increase of global

---

<sup>3</sup>See Finus and Caparrós (2015) for a recent survey of this literature and a collection of the most influential articles since the early papers by Barrett (1994a) and Carraro and Siniscalco (1993).

<sup>4</sup>Intuitively, when the gains from cooperation are large, cooperation requires a major departure from non-cooperative behavior, associated with large free-rider incentives and vice versa.

welfare - BCAs may impose serious welfare costs as long as not full cooperation is achieved.

Secondly, we are interested how different membership rules affect the success of BCAs. We show that signatories may find it attractive to restrict accession to their agreement before full participation is reached due to strategic reasons. Consequently, under an exclusive membership rule, a fully cooperative agreement may not be obtained, which may go along with large welfare costs at the world scale. We analyze this issue by considering a simultaneous and sequential process of coalition formation.

This paper is an extension of Finus and Al Khourdajie (2017), which considers taxation in an intra-industry trade model with horizontal product differentiation and consumers' taste for variety (TFV) and coalition formation, by adding import tariffs. We do not consider OBRs for two reasons. Firstly because we can show that BCAs even without OBRs can lead to successful climate agreements, in particular if they are in the tradition of environmental treaties of the open membership type. Secondly, OBRs would undermine the strong polluter principle. Domestic production is harmful irrespective of the market to which goods are sold and even with OBRs, the harmful production of non-signatories for their own market are not internalized. Hence, a justification within WTO-rules seems much weaker. We briefly discuss our and alternative assumptions in the concluding Section 6.

Our model benefits from three strands of literature. Firstly, the literature on strategic trade models, which extended the model by Brander and Spencer (1985) by including environmental damages and consumer surplus in governments welfare function (e.g. Barrett 1994b, Conrad 1993, Kennedy 1994 and Ulph 1996), though this literature has mainly restricted attention to only two countries and/or ignored the issue of agreement formation. Secondly, the literature on international environmental agreements, which focuses on treaty formation and emission reduction, but typically ignores trade.<sup>5</sup> Thirdly, the literature on trade agreements, though without considering environmental issues, in a very similar setting than ours and from which it emerges that membership rules may determine the success of treaty formation (e.g. Yi 1996 and 2000).

Our paper proceeds as follows. In section 2, we present the three-stage coalition formation model in which countries choose first their membership, then choose their policy levels and finally firms choose their output. We solve the game by backwards induction and hence consider the output stage in section 3, the policy stage in section 4 and the membership stage in section 5, including an overall evaluation of stable agreements. Section 6 concludes, qualifies our results and points to future research.

---

<sup>5</sup>See the literature cited in Finus and Caparrós (2015). Eichner and Pethig (2013, 2014, 2015) have recently introduced trade in the IEA-literature, though their model is very different from our intra-industry trade model and they do not consider BCAs.

## 2 Model

### 2.1 Coalition Formation Game

Consider an intra-industry trade model with  $n$  *ex-ante* symmetric countries with a representative firm and consumer in each country. We denote the set of countries by  $N$  with  $n$  the number of countries, the cardinality of  $N$ . Firms produce a horizontally differentiated good, i.e. the same good but in different varieties where each firm produces one variety. This good causes environmental damages; the production uses emissions as an input, for example in the form of energy by which greenhouse gases are released. Firms compete in a Nash-Cournot fashion. Markets are segmented and each firm supplies its good to the domestic and all foreign markets. Because of the segmentation of markets, firms play a separate Cournot-game in each market. Transport costs are assumed away as usual.<sup>6</sup>

We assume a three-stage coalition formation game:

*Stage 1, Choice of Membership:* all countries decide simultaneously whether to join coalition  $S \subseteq N$  with  $m$  the cardinality of  $S$ ,  $1 \leq m \leq n$ . Countries which do not join  $S$  act as singletons. A typical signatory will be denoted by  $i$  and a non-signatory by  $j$ .

*Stage 2, Choice of Policy Level:* all countries choose simultaneously their emission tax.

- Signatories choose their tax  $t_i$  (implemented uniformly in all signatory countries) by maximizing the joint welfare of coalition  $S$ :  $\max_{t_i} \sum_{i \in S} W_i$ .
- Non-signatories choose their individual tax  $t_j$  by maximizing their individual welfare:  $\max_{t_j} W_j$ .

*Stage 3, Choice of Output:* all firms choose simultaneously and non-cooperatively their segmented market outputs by maximizing their producer surplus:  $\max_{q_{1i}, \dots, q_{ni}} PS_i$ .

The game is solved by backwards induction.

In the third stage, firms play a Nash equilibrium output game in each segmented market, facing taxes. Output  $q_{ki}$  refers to the output of firm  $i$  sold in market  $k$ . We consider two tax regimes. The first regime is called *No-BCA-regime*. Each government imposes a tax on its representative firm. Firms in signatory countries face tax  $t_i$  and firms in non-signatory countries face tax  $t_j$ . As will become clear below, the emissions tax is equivalent to an output tax as the tax is imposed per unit of output and we assume a constant emission output ratio which we normalize

---

<sup>6</sup>Thus, our model is in the spirit of for instance Loke and Winter (2012) and Yi (1996, 2000). See also Helpman and Krugman (1985).



to one. The second regime is called the *BCA-regime*. It is almost identical to the first regime, except for firms in non-signatory countries. For quantities sold to their own market or any other non-signatory market, they again face tax  $t_j$ . Also for quantities sold to signatory markets, they need to pay tax  $t_j$  as usual if  $t_j \geq t_i$ . However, if  $t_j < t_i$ , they need to pay additionally a mark-up  $\phi(t_i - t_j)$ , i.e. the border carbon adjustment and hence they pay  $t_j + \phi(t_i - t_j)$  per unit of output. Following Eyland and Zaccour (2013), we call  $\phi$  the BCA-adjustment parameter. In accordance with the anti-discrimination rules of the WTO,  $\phi > 1$  is not feasible and hence we need to assume  $0 \leq \phi \leq 1$ . Obviously, for  $\phi = 0$ , BCA- and No-BCA-regime coincide. If signatories can use border carbon adjustments in stage 2, it seems sensible to assume  $\phi = 1$ , not only as a benchmark case, but also because signatories would have no reason to choose any smaller value could they choose this parameter optimally.<sup>7</sup> This equips signatories with the “maximum possible enforcement power”. The third stage delivers an equilibrium vector of outputs of firms located in signatory and non-signatory countries which is a function of all taxes.

In the second stage, inserting equilibrium outputs from stage 3 in countries’ welfare function, we can solve for countries equilibrium tax rates. Signatories choose their taxes by acting as one single player whereas each non-signatory acts as a single player. Hence, the solution of the second stage can be interpreted as a Nash equilibrium tax game between coalition  $S$  with  $m$  players and  $n - m$  singleton players. As we will see, equilibrium taxes depend on the size of coalitions and whether border carbon adjustments are available to signatories. If we insert equilibrium tax rates (stage 2) into equilibrium outputs (stage 3) and those in welfare functions, welfare of signatories and non-signatories can be expressed as a function of coalition  $S$  only,  $W_i(S)$  and  $W_j(S)$ , respectively. This provides the input for the first stage.

In the first stage, we solve for a Nash equilibrium in membership strategies in a cartel formation game (d’Aspremont et al. 1983), also called a simultaneous move open-membership single coalition game (Yi 1997) in order to stress the institutional setting of this type of agreement. Each country simultaneously chooses whether to join coalition  $S$  or to remain a singleton. The treaty is of the open-membership type because nobody can be excluded from joining coalition  $S$ . A coalition is called stable if no player has an incentive to change her announcement, given the announcement of all other players. Following d’Aspremont et al. (1983), we can also say that no signatory has an incentive to leave coalition  $S$  (internally stability:  $W_i(S) - W_i(S \setminus \{i\}) \geq 0 \forall i \in S$ ) and no non-signatory has an incentive to join coalition  $S$  (external stability:  $W_j(S) - W_j(S \cup \{j\}) \geq 0 \forall j \in N \setminus S$ ). In the course of the later discussion, we will also consider and introduce alternative institutional settings, e.g. a sequential choice of membership and exclusive membership.

We note that if the grand coalition forms (a coalition of all players), this replicates the social optimum, and if no coalition forms, i.e. all players act as singletons, this

---

<sup>7</sup>It can be shown that signatory countries would choose  $\phi > 1$  if  $\phi$  was treated as a choice variable.

replicates a Nash equilibrium in a game without coalition formation. Moreover, if the grand coalition forms, and hence there is no outsider left in the game, the taxes under the BCA- and No-BCA-regime coincide, which are socially optimal. Also if all players act as singletons, so no agreement has formed, no tariff at the border can be implemented, and, again, taxes under both regimes coincide.

## 2.2 Welfare Function

In this subsection, we have a closer look at the welfare function of governments and its different components.

If a country  $i$  becomes a signatory,  $i \in S$ , its welfare is given by:

$$W_i = CS_i + PS_i - D_i + TR_i + BCR_i \quad (1)$$

where  $CS_i$  represents country  $i$ 's consumer surplus,  $PS_i$  country  $i$ 's producer surplus,  $D_i$  is the pollution damage faced by country  $i$ ,  $TR_i$  is country  $i$ 's tax revenue from the tax imposed by the government on its domestic firm's production, and  $BCR_i$  is country  $i$ 's tariff from the border carbon adjustment imposed by the domestic government on imports from firms located in non-signatory countries. Of course, under the No-BCA-regime  $BCR_i = 0$ .

If a country  $j$  remains a non-signatory,  $j \in N \setminus S$ , its welfare is given by:

$$W_j = CS_j + PS_j - D_j + TR_j \quad (2)$$

with the same welfare components as in equation (1) after the appropriate changes of subscripts. Note that the BCR-term is missing in a non-signatory government's welfare function as per assumption they cannot impose a tariff on imports.

Consumers in all countries have identical preferences which are represented by a quasi-linear utility function over two goods:

$$u_i(q_i; M_i) = v_i(q_i) + M_i = aQ_i - \frac{\gamma}{2}Q_i^2 - \frac{1-\gamma}{2} \sum_{k \in N} q_{ik}^2 + M_i \quad (3)$$

where  $v_i$  represents the utility from consuming the horizontally differentiated and traded good and  $M_i$  represents the utility from consuming the numeraire good, representing the composition of all other goods;  $q_i = (q_{i1}, \dots, q_{in})$  is a vector of the varieties consumed by consumers in country  $i$  that are produced by all signatories' and non-signatories' firms, with  $q_{ik}$  representing country  $i$ 's consumption of country  $k$ 's variety<sup>8</sup>;  $a$  is a positive demand parameter and  $Q_i = \sum_{k \in N} q_{ik}$  is country  $i$ 's total

---

<sup>8</sup>Throughout the paper the first subscript indicates the market in which the variety is consumed and the second subscript indicates the country in which it is produced.

consumption of all varieties, supplied by all firms  $k$  (i.e. located in signatory and non-signatory countries). Thus, utility is linear in the numeraire good and quadratic in the differentiated good. Consumers have a taste for variety (Dixit and Stiglitz, 1977). Hence, their utility depends not only on the total quantity consumed (term  $aQ_i - \frac{\gamma}{2}Q_i^2$  in (3)) but also on the composition of quantities of the differentiated good (term  $-\frac{1-\gamma}{2}\sum_{k \in N} q_{ik}^2$  in (3)) as for instance assumed in Yi (1996 and 2000). The taste for variety (abbreviated TFV hereafter) is captured by parameter  $\gamma \in [0, 1]$ . High values of  $\gamma$  imply a low taste for variety and low values of  $\gamma$  correspond to a high taste for variety. For  $\gamma = 1$  varieties are perfect substitutes and for  $\gamma = 0$  varieties cannot be substituted at all.<sup>9</sup>

From (3), consumer  $i$ 's inverse demand function for country  $k$ 's variety follows from:

$$p_{ik} = \frac{\partial u_i}{\partial q_{ik}} \iff p_{ik} = a - (1 - \gamma)q_{ik} - \gamma Q_i. \iff p_{ik} = a - q_{ik} - \gamma \sum_{l \in N, l \neq k} q_{il} \quad (4)$$

where  $p_{ik}$  represents the price faced by consumers in country  $i$  consuming the variety produced by a firm located in country  $k$  and  $\sum_{l \in N, l \neq k} q_{il}$  is the sum of all consumed varieties produced by all firms except the firm located in country  $k$ .

From (3) and (4), the consumer surplus in country  $i$  is given by:

$$CS_i = aQ_i - \frac{\gamma}{2}Q_i^2 - \frac{1-\gamma}{2}\sum_{k \in N} q_{ik}^2 - \sum_{k \in N} q_{ik}p_{ik} \quad (5)$$

where the last term in (5) is the representative consumer  $i$ 's spending.

For producers, allowing for the possibility of border carbon adjustments, we need to distinguish between firms located in signatory and non-signatory countries. The producer surplus of a firm located in a signatory country  $i$  is the sum of its profit obtained in each market:

$$PS_i = \sum_{k \in S} \pi_{ki} + \sum_{l \in N \setminus S} \pi_{li} = \sum_{k \in S} q_{ki}(p_{ki} - c - t_i) + \sum_{l \in N \setminus S} q_{li}(p_{li} - c - t_i) \quad (6)$$

where  $\pi_{ki}$  ( $\pi_{li}$ ) represents firm  $i$ 's profit in signatory  $k$ 's (non-signatory  $l$ 's) market from selling quantity  $q_{ki}$  ( $q_{li}$ ) at price  $p_{ki}$  ( $p_{li}$ );  $c$  is a constant marginal cost parameter and  $t_i$  is the tax imposed by country  $i$ 's government on its firm's production.

Also the producer surplus of a firm located in a non-signatory country  $j$  is the sum of its profit in each market:

$$PS_j = \sum_{k \in S} \pi_{kj} + \sum_{l \in N \setminus S} \pi_{lj} = \sum_{k \in S} q_{kj}(p_{kj} - c - t_j - \Omega) + \sum_{l \in N \setminus S} q_{lj}(p_{lj} - c - t_j) \quad (7)$$

---

<sup>9</sup>An extension could be the "ideal variety" approach where consumers have not only a general preference for the variety of a good but also a preference for a particular variety. One application is a preference for the domestically produced variety (Di Comite et al. 2014).

$$\text{with } \Omega = \begin{cases} \phi(t_i - t_j) & \text{if } t_i > t_j \\ 0 & \text{if } t_i \leq t_j \end{cases}$$

where  $\pi_{kj}$  ( $\pi_{lj}$ ) represents firm  $j$ 's profit in signatory  $k$ 's (non-signatory  $l$ 's) market from selling quantity  $q_{kj}$  ( $q_{lj}$ ) at price  $p_{kj}$  ( $p_{lj}$ ), and  $t_j$  is the tax imposed by country  $j$ 's government on its firm's production. Clearly, under the No-BCA-regime,  $\phi = 0$ , therefore  $\Omega = 0$  and hence the structure of the producer surplus of a firm located in a signatory and non-signatory country is the same. Under the BCA-regime, this is also the case if  $t_i \leq t_j$  because then  $\Omega = 0$ , but is different if  $t_i > t_j$  because then  $\Omega > 0$  provided there is some adjustment, i.e.  $\phi > 0$ . As mentioned above, we assume full adjustment under the BCA-regime and hence set  $\phi = 1$  later on. Then  $t_j + \Omega = t_i$ , i.e. firms located in non-signatory countries face de facto signatories' tax  $t_i$  for all quantities which they export to signatories' markets. For all other quantities they face only their local tax  $t_j$ .

Damages in signatory and non-signatory countries are the same ( $D_i = D_j = D_l$ ) and are given by:

$$D_l = \delta Q \quad (8)$$

where  $\delta$  is a damage parameter,  $Q = \sum_{i \in N} Q_i = \sum_{i \in N} Q_i$  is total production which is equal to total consumption world-wide. That is, we assume a constant emission output coefficient which we normalize to 1. Hence, we assume that emissions constitute a pure public bad: damages depend on total emissions. As there is no abatement technology in our simple model, emission and output tax are the same and an output/emission tax is an efficient policy instrument to address externalities.

The tax revenue of signatory country  $i$  is given by:

$$TR_i = t_i \sum_{k \in N} q_{ki} \quad (9)$$

which is the tax rate it imposes on its firm multiplied by the total quantity produced by its firm for all markets. Similarly, the tax revenue of non-signatory country  $j$  is given by:

$$TR_j = t_j \sum_{k \in N} q_{kj} . \quad (10)$$

Finally, under the BCA-regime, the border carbon adjustment revenue of a signatory country obtained from the adjustment  $\Omega$  on imports from firms located in non-signatory countries is given by:

$$BCR_i = \Omega \sum_{j \in N \setminus S} q_{ij} \quad (11)$$

with  $\Omega$  defined in equation (7).

### 3 Third Stage

In this section, we derive output of firms in the third stage. The background of this derivation is provided in Appendix 1 which draws on the description of the coalition formation game in subsection 2.1 and the description of the welfare components in subsection 2.2. Subsection 2.1 is relevant in that  $m$  signatories jointly choose a uniform tax  $t_i$  and  $n - m$  non-signatories choose individually their tax  $t_j$ . In the following, for simplicity, we already use the information from stage 2 that, in equilibrium, not only all signatories will choose the same  $t_i$  but also all non-signatories choose the same tax  $t_j$  (because of identical welfare functions), though normally  $t_i \neq t_j$ . From subsection 2.2 we use the demand function in each market and the definition of the producer surplus, which firms maximize by choosing their output vector, i.e. the quantities of their variety they sell to different markets. This output vector has the structure displayed in the table below.

	Sign. firm $i$	Non-sign. firm $j$	Total Consumption
Sign. market $k$	$q_{ki}^*$	$q_{kj}^*$	$Q_{k.}^*$
Non-sign. market $l$	$q_{li}^*$	$q_{lj}^*$	$Q_{l.}^*$
Total Production	$Q_{.i}^*$	$Q_{.j}^*$	$Q$

In the following, we start by listing the quantities consumed in different markets and then display aggregate production levels. In order to cover both regimes, the No-BCA- and the BCA-regime, we display quantities with the adjustment parameter  $\phi$  where appropriate, noting that by setting  $\phi = 0$  ( $\phi = 1$ ) quantities in the No-BCA- (BCA-) regime would be obtained. It is important to note that quantities are a function of equilibrium taxes, which are determined in stage 2 and those taxes will differ between the No-BCA- and the BCA-regime. Hence for instance, equilibrium quantity  $q_{li}^*$  in (12) below will differ in both regimes because equilibrium taxes differ. By the same token, this applies to all quantities listed below.

A signatory firm  $i$ 's variety produced for a non-signatory  $l$ 's market is given by

$$q_{li}^* = \frac{(a - c)(2 - \gamma) - [\gamma(n - m) + (2 - \gamma)]t_i + [\gamma(n - m)]t_j}{((n - 1)\gamma + 2)(2 - \gamma)} \quad (12)$$

and a non-signatory firm  $j$ 's variety produced for a non-signatory  $l$ 's market is

$$q_{lj}^* = \frac{(a - c)(2 - \gamma) + \gamma m t_i - [2 + \gamma(m - 1)]t_j}{((n - 1)\gamma + 2)(2 - \gamma)} \quad (13)$$

with the total equilibrium consumption in a non-signatory  $l$ 's market from all varieties given by:

$$Q_{l.}^* = \frac{n(a - c) - m t_i - (n - m)t_j}{(n - 1)\gamma + 2}. \quad (14)$$

Hence, in non-signatory markets, the adjustment parameter  $\phi$  does not play a role. This is different in signatories' markets.

A signatory firm  $i$ 's variety produced for a signatory  $k$ 's market is given by:

$$q_{ki}^* = \frac{(a-c)(2-\gamma) - [\gamma(n-m)(1-\phi) + (2-\gamma)] t_i + [\gamma(n-m)(1-\phi)] t_j}{((n-1)\gamma + 2)(2-\gamma)} \quad (15)$$

and a non-signatory firm  $j$ 's variety produced for a signatory  $k$ 's market is given by:

$$q_{kj}^* = \frac{(a-c)(2-\gamma) + [\gamma m - \phi(2 + \gamma(m-1))] t_i - [2 + \gamma(m-1) - \phi(2 + \gamma(m-1))] t_j}{((n-1)\gamma + 2)(2-\gamma)} \quad (16)$$

with the total equilibrium consumption in a signatory  $k$ 's market from all varieties given by:

$$Q_k^* = \frac{n(a-c) - [n - (n-m)(1-\phi)] t_i - [(n-m)(1-\phi)] t_j}{(n-1)\gamma + 2}. \quad (17)$$

We now look at total production. Total production of a signatory's firm  $i$ 's variety to all markets is given by:

$$Q_{.i}^* = \frac{n(a-c)(2-\gamma) - [2n + \gamma n(n-1-m) - \phi \gamma m(n-m)] t_i}{((n-1)\gamma + 2)(2-\gamma)} + \frac{[n(n-m) - \phi m(n-m)] t_j}{((n-1)\gamma + 2)(2-\gamma)} \quad (18)$$

and total production of a non-signatory's firm  $j$ 's variety to all markets is given by:

$$Q_{.j}^* = \frac{n(a-c)(2-\gamma) + [\gamma nm - \phi m(2 + \gamma(m-1))] t_i^*}{((n-1)\gamma + 2)(2-\gamma)} - \frac{[\gamma nm + n(2-\gamma) - \phi m(2 + \gamma(m-1))] t_j^*}{((n-1)\gamma + 2)(2-\gamma)} \quad (19)$$

and finally, total production/consumption by all countries is given by:

$$Q = \frac{n^2(a-c) - [nm + \phi m(n-m)] t_i^* - [n(n-m) - \phi m(n-m)] t_j^*}{(n-1)\gamma + 2}. \quad (20)$$

In order to render the comparison between the two regimes interesting, henceforth, when referring to the BCA-regime, we assume  $t_i > t_j$ , i.e. signatories choose a higher tax (or lower subsidy) than non-signatories. Using the quantities listed in (12) to (20), some basic algebra allows to draw the following conclusions.

### Proposition 1 - The Effect of Taxes on Equilibrium Quantities

Suppose a coalition  $S$  with  $m$  signatories has formed in the first stage, and let  $1 < m < n$ .

- Under the No-BCA-regime, the quantity of firm  $i$ 's ( $j$ 's) variety supplied in any market  $h \in N$  decreases in signatories' (non-signatories') taxes,  $\frac{\partial q_{hi}^*}{\partial t_i} < 0$  ( $\frac{\partial q_{hj}^*}{\partial t_j} < 0$ ), and increases in non-signatories' (signatories') taxes,  $\frac{\partial q_{hi}^*}{\partial t_i} > 0$  ( $\frac{\partial q_{hj}^*}{\partial t_j} > 0$ ), except for  $\gamma = 0$  in which case  $\frac{\partial q_{hi}^*}{\partial t_j} = 0$  ( $\frac{\partial q_{hj}^*}{\partial t_i} = 0$ ). Under the BCA-regime, the same is true for all quantities supplied to non-signatories' markets  $l \in N \setminus S$ , i.e.  $\frac{\partial q_{li}^*}{\partial t_i} < 0$  ( $\frac{\partial q_{lj}^*}{\partial t_j} < 0$ ) and  $\frac{\partial q_{li}^*}{\partial t_j} > 0$  ( $\frac{\partial q_{lj}^*}{\partial t_i} > 0$ ), except for  $\gamma = 0$  in which case  $\frac{\partial q_{li}^*}{\partial t_j} = 0$  ( $\frac{\partial q_{lj}^*}{\partial t_i} = 0$ ). For quantities supplied to all signatories' markets  $k \in S$ , the quantities of firms  $i$ 's and  $j$ 's varieties decrease in signatories' taxes,  $\frac{\partial q_{ki}^*}{\partial t_i} < 0$  and  $\frac{\partial q_{kj}^*}{\partial t_i} < 0$ , whereas they remain unaffected by non-signatories' taxes,  $\frac{\partial q_{ki}^*}{\partial t_j} = 0$  and  $\frac{\partial q_{kj}^*}{\partial t_j} = 0$ .
- Under the No-BCA-regime, the total quantity produced by firm  $i$  ( $j$ ) decreases (increases) in signatories' taxes and increases (decreases) in non-signatories' taxes,  $\frac{\partial Q_{.i}^*}{\partial t_i} < 0$  and  $\frac{\partial Q_{.i}^*}{\partial t_j} > 0$  ( $\frac{\partial Q_{.j}^*}{\partial t_i} > 0$  and  $\frac{\partial Q_{.j}^*}{\partial t_j} < 0$ ), except for  $\gamma = 0$  in which case  $\frac{\partial Q_{.i}^*}{\partial t_j} = 0$  ( $\frac{\partial Q_{.j}^*}{\partial t_i} = 0$ ). Under the BCA-regime, the same is true for all firms  $i$  located in signatory countries, i.e.  $\frac{\partial Q_{.i}^*}{\partial t_i} < 0$  and  $\frac{\partial Q_{.i}^*}{\partial t_j} > 0$ , except for  $\gamma = 0$  in which case  $\frac{\partial Q_{.i}^*}{\partial t_j} = 0$ . For firms  $j$  located in non-signatory countries  $\frac{\partial Q_{.j}^*}{\partial t_j} < 0$  but  $\frac{\partial Q_{.j}^*}{\partial t_i} > 0$  if and only if  $\gamma(n - m + 1) - 2 > 0$  (otherwise  $\frac{\partial Q_{.j}^*}{\partial t_i} \leq 0$ ).
- Under the No-BCA-regime, the total quantity consumed in any market  $h \in N$  decreases in both taxes,  $\frac{\partial Q_{h.}^*}{\partial t_i} < 0$  and  $\frac{\partial Q_{h.}^*}{\partial t_j} < 0$ . Under the BCA-regime, the same is true for all quantities supplied to all non-signatories' markets  $l \in N \setminus S$ , i.e.  $\frac{\partial Q_{l.}^*}{\partial t_i} < 0$  and  $\frac{\partial Q_{l.}^*}{\partial t_j} < 0$ . For quantities supplied to all signatories' markets  $k \in S$ , the total quantity consumed in a signatory's market decreases in signatories' taxes but are unaffected by non-signatories' taxes:  $\frac{\partial Q_{k.}^*}{\partial t_i} < 0$  and  $\frac{\partial Q_{k.}^*}{\partial t_j} = 0$ .

**Proof:** Follows directly from equations (12) to (20) above, assuming  $\phi = 0$  under the No-BCA- and  $\phi = 1$  and  $t_i > t_j$  under the BCA-regime. **Q.E.D.**

Under the No-BCA-regime, as expected, quantities of any firm's variety (signatory's or non-signatory's firm) produced for any market (signatories' and non-signatories' markets) are negatively affected by domestic taxes and positively affected by foreign taxes. Only for full TFV, i.e.  $\gamma = 0$ , will a firm's output not be (positively) affected by foreign taxes (imposed on foreign firms). In this particular case, each firm acts

like a monopolist in each market as consumers do not substitute between varieties and hence there is no competition among firms. Under the BCA-regime, the same strategic interaction applies in non-signatories' markets. However, in signatories markets, all firms face de facto the same tax  $t_i$  under BCAs and therefore non-signatories' taxes  $t_j$  do not affect quantities supplied in these markets.

This strategic interaction between domestic and foreign taxes is also reflected at the level of total production. Total quantities produced by a firm are negatively affected by domestic taxes and positively by foreign taxes; only for full TFV, production levels are not affected by foreign taxes. Under the BCA-regime, there is one exception from this general pattern. The total production of a firm located in a non-signatory country may be negatively affected by signatories' taxes as there are two effects. Quantities sold to non-signatories markets are positively affected by signatories' taxes (as usual; see (13)), but quantities sold to signatories' markets are negatively affected (see 16). Overall, if the value of  $\gamma$  is small and/or the coalition is large (large value of  $m$ ), the negative effect dominates the positive effect.

Finally, total consumption in every country decreases in signatories' and non-signatories' taxes. Only under the BCA-regime, the consumption in a signatory country is not affected by non-signatories' taxes.

Taken together, under the BCA-regime, governments in signatory countries jointly enjoy more market power by protecting their domestic markets through border carbon adjustments than under the No-BCA-regime. That is, if signatory governments choose a higher tax than non-signatory governments under the BCA-regime, at least in signatories' markets all firms face de facto the same tax  $t_i$  and hence act on an equal playing field. The following proposition sheds further light on the difference of production and consumption patterns under the two regimes. In order to focus the analysis, henceforth, we assume  $t_i > t_j$  not only under the BCA- but also under No-BCA-regime. In section 4, we will discuss this assumption in more detail where we derive equilibrium taxes in the second stage of the game.

## **Proposition 2 - The Effects of Taxes on Production and Consumption Patterns**

*Suppose a coalition  $S$  with  $m$  signatories has formed in the first stage, and let  $1 < m < n$ . Moreover, assume that governments have chosen their taxes in the second stage and let  $t_i > t_j$  under the No-BCA- and BCA-regime.*

- *Under the No-BCA-regime, signatory firm  $i$ 's outputs (non-signatory firm  $j$ 's outputs) for all markets are the same,  $q_{ki}^* = q_{li}^*$  ( $q_{lj}^* = q_{kj}^*$ ). Under the BCA-regime, outputs are differentiated. Signatory firm  $i$ 's outputs for signatories' markets  $k \in S$  are higher than their outputs for non-signatories' markets  $l \in S \setminus N$ ,  $q_{ki}^* > q_{li}^*$ , except for  $\gamma = 0$  in which case  $q_{ki}^* = q_{li}^*$ . Non-signatory firm*



*j's outputs for non-signatories' markets  $l \in S \setminus N$  are higher than their outputs for signatories' markets  $k \in S$ ,  $q_{lj}^* > q_{kj}^*$ .*

- *In every signatory market  $k \in S$ , under the No-BCA-regime, signatory  $i$ 's output is lower than non-signatory firm  $j$ 's output,  $q_{ki}^* \leq q_{kj}^*$  whereas under the BCA-regime, signatory  $i$ 's output and non-signatory firm  $j$ 's output are the same,  $q_{ki}^* = q_{kj}^*$ .*
- *Under both regimes, in every non-signatory market  $l \in S \setminus N$ , a non-signatory firm  $j$ 's output is always higher than a signatory firm  $i$ 's output,  $q_{li}^* < q_{lj}^*$ .*
- *Consequently, total production of a signatory firm  $i$  is lower than of a non-signatory firm  $j$  under both regimes,  $Q_i^* < Q_j^*$ . Moreover, total consumption in all markets is the same under the No-BCA-regime,  $Q_k^* = Q_l^*$ , but is lower in every signatory market  $k \in S$  than in every non-signatory market  $l \in S \setminus N$  under the BCA-regime,  $Q_k^* < Q_l^*$ .*

**Proof:** Follows directly from equations (12) to (20) above, assuming  $t_i > t_j$ ,  $\phi = 0$  under the No-BCA- and  $\phi = 1$  under the BCA-regime. **Q.E.D.**

Whereas under the No-BCA-regime, firms supply the same quantities to all markets, they differentiate quantities under the BCA-regime. Under the No-BCA-regime only taxes but not markets matter for firms' profits. In contrast, under the BCA-regime, markets matter because taxes are different in different markets. For firms located in signatory countries, signatories' markets are more attractive because also their competitors face tax  $t_i$  whereas in non-signatories' markets they also face  $t_i$  but their competitors only  $t_j$ . For firms located in non-signatory countries, the mirror image argument applies and hence they prefer to sell more to non-signatories' than signatories' markets.

Generally speaking, quantities sold in each market depend on relative taxes. Because taxes in signatory countries are higher than in non-signatory countries by assumption, firms located in non-signatory countries sell more in each market and hence have higher profits than firms located in signatory countries. This advantage of non-signatory over signatory firms is partially offset through BCAs because quantities and hence profits in signatory markets are now the same (though quantities and profits in non-signatory markets are still higher). Not surprisingly, at the aggregate production level, under both regimes, non-signatory firms produce more than signatory firms and hence earn higher profits. Thus, taken together, BCAs allow to reduce differences in profits between signatory and non-signatory firms but cannot make this difference disappear.

It is important to note that this conclusion does not change when moving from firms' perspective to an aggregate welfare perspective. Signatory firms' profits are lower than non-signatory firms' profits for two reasons. Lower output and higher taxes. However from a countries welfare perspective taxes are welfare neutral in this simple

model: firms' tax bills are revenues of governments. Thus, what matters at the aggregate are profits excluding tax payments, i.e. gross profits. Hence, also from this aggregate welfare perspective, signatory countries are disadvantaged because of lower gross profits, which is a result of lower output.<sup>10</sup>

Regarding consumption, under the No-BCA-regime, consumers in signatory and non-signatory countries consume exactly the same individual and hence total quantities. Hence, the consumer surplus in signatory and non-signatory countries is the same. In contrast, under the BCA-regime, total consumption in signatories' markets is lower than in non-signatories' markets as all quantities supplied to signatory markets face the high tax  $t_i$  whereas those supplied to non-signatories' markets depend not only on  $t_i$  but also on lower  $t_j$ . This also implies that the quantity for each variety consumed is lower in a signatory market than in a non-signatory market, and, consequently, also the consumer surplus is lower in signatory than in non-signatory countries. Thus, it is important to note that the border carbon adjustment does not improve but worsens the situation compared to the No-BCA-regime.

In terms of environmental damages, even though total production and hence global pollution may be reduced when moving from the No-BCA- to the BCA-regime, in relative terms nothing changes between signatory and non-signatory countries, as damages are the same in both types of countries.

Finally, there is an additional factor that comes into play when moving from the No-BCA- to the BCA-regime, the revenues from the border carbon adjustment. This is de facto a transfer from non-signatory to signatory countries, respectively from the firms located in non-signatory countries to signatory governments.

Hence, overall, implementing a BCA-regime implies reducing the competitive advantage of non-signatory over signatory firms and hence reducing the difference in gross profits, collecting revenues from foreign firms by signatory governments, but disadvantaging consumers in signatory countries. As our later analysis will illustrate, the overall effect works to the advantage of signatory over non-signatory countries, explaining, among other factors, why larger agreements are stable under the BCA-compared to the No-BCA-regime, though it will turn out that we need to be careful in rushing to conclusions regarding global welfare.

## 4 Second Stage

In the second stage, governments choose their taxes. All governments understand how taxes affect quantities and how they affect welfare. That is, governments have solved the third stage of the game. As pointed out in subsection 2.1, the difference is that governments which are part of coalition  $S$ , choose their tax  $t_i$  by maximizing the aggregate welfare of all coalition members, and governments which do not belong

---

<sup>10</sup>A proof is straightforward and available upon request.

to  $S$  choose their tax  $t_j$  by maximizing their individual welfare. Taxes of other governments are taken as given.<sup>11</sup> The simultaneous solution of the  $m$  first order conditions of signatory governments and  $n - m$  first order conditions of non-signatory governments delivers equilibrium taxes, which are a function of all parameters of the model, but in particular they depend on the size of coalition  $S$ ,  $m$ , and hence we may write  $t_i^*(m)$  and  $t_j^*(m)$ . It turns out that welfare functions are strictly convex in own taxes,  $\partial^2 W_{i \in S}(t_i)/\partial t_i^2 < 0$  and  $\partial^2 W_{j \notin S}(t_j)/\partial t_j^2 < 0$ , and that the Hessian matrix is semi-definite, guaranteeing a unique stable equilibrium tax vector.

Inserting equilibrium taxes in equilibrium quantities as displayed in equations (12) to (20) in Section 3, reveals that we need to impose constraints on the parameters such that equilibrium quantities are positive. Such non-negativity constraints essentially boil down to the condition that the demand parameter  $a$  is larger than marginal production cost  $c$  plus a multiple of marginal damages  $\delta$ . In other words, these non-negativity constraints represents a lower threshold  $\underline{a}$ , such that if  $a \geq \underline{a}$  holds, all quantities are positive for all  $m$ ,  $1 \leq m \leq n$ . Typically, the most restrictive condition applies in the grand coalition ( $m = n$ ) with the highest overall tax level. For instance, under the No-BCA-regime, the non-negativity constraint is  $a \geq c + m\delta$  for  $\gamma = 0$  and  $a \geq c + n\delta\Psi_A/m\Psi_B$  with  $\Psi_A = n^2(m - 1) - n(m - 1)^2 - m(m - 2)$ ,  $\Psi_B = n^2 - (m - 1)(n + 1)$ ,  $\Psi_A > 0$  and  $\Psi_B > 0$  for  $\gamma = 1$ . For  $\gamma = 0$  it is immediately evident that the condition is most restrictive if  $m = n$ . However for  $\gamma = 1$ , this is also the case because  $\Psi_A/m\Psi_B$  increases in  $m$ . By setting  $m = n$ , the non-negativity constraint is given by  $a > c + n\delta$  for both values of  $\gamma$ . Clearly, an alternative interpretation of this condition is that global marginal damages cannot be too high compared to the difference  $a - c$ , which can be interpreted as the market size corrected for production costs. For the BCA-regime, similar non-negativity constraints can be derived, though they look much more complicated, but, in essence, they also require that global marginal damages cannot be too high compared to the “corrected” market size.

An other issue which is important for our analysis is that we want to consider situations for which  $t_i^* \geq t_j^*$  holds. In our model (like in many other strategic trade models), this is not automatically guaranteed. The reason is simple: in this game, we have two market imperfections. On the one hand, there is Cournot-competition in international trade. This implies that signatories, which internalize externalities among its members, have an incentive to subsidize their consumers. However, they also have an incentive to tax their producers in order to enforce a cartel solution, i.e. stabilizing the market price by reducing output and output is reduced through taxes. Different from models with only two countries and no coalition formation, in our model, signatory governments have an incentive to tax their producers. The

---

<sup>11</sup>Technically, quantities as a function of taxes as known from stage three are inserted in governments’ welfare functions and the welfare functions are differentiated with respect to own taxes. Importantly, different from the quantities displayed in section 3, which already used the information of symmetry, quantities must be expressed as function of the full tax vector, and the symmetry assumption can only be invoked after derivatives have been taken.

reason is simple. In our model, taxes, which producers have to pay, are welfare neutral as they constitute revenues for the government. The gross profits of firms (i.e profits excluding tax bills) are maximized in a monopoly. As firms compete non-cooperatively in each market as long as  $\gamma > 0$ , signatory governments aiming to maximize the gross profits of their firms have an incentive to collude. Trading off firm's gross profits for consumer surplus, leads to a subsidy if the market size is large compared to production cost, represented by the parameters  $a$  and  $c$  in our model. On the other hand, there is global pollution, which calls for taxes in order to reduce damages. The importance of damages in the welfare function is represented by the parameter  $\delta$  in our model. Hence, as known from many strategic trade models, even if the grand coalition forms, taxes may not be set at the Pigouvian level, i.e. not equal to the sum of marginal damages  $n\delta$ . They may be set lower or higher, depending on the importance consumers and producers receive in governments' welfare function. Thus, if the value of the parameter  $a$  is high compared to the value of the parameters  $c$  and  $\delta$ , signatory governments may chose a lower tax than non-signatory governments and taxes may even be negative, i.e. governments subsidize their firms. For instance, under the No-BTA-regime and  $\gamma = 0$ ,  $t_i^*(m) \geq t_j^*(m) > 0$  if  $a \leq c + 2n\delta$  but  $t_i^*(m) < t_j^*(m) < 0$  if  $a > c + 2n\delta$ , both inequalities are compatible with the non-negativity constraint derived above, i.e.  $a \geq c + n\delta$ . For  $\gamma = 1$ , one can show that  $t_i^*(m) > t_j^*(m)$  and  $t_j^*(m) < 0$  always hold whereas  $t_i^*(m)$  can be positive or negative without violating the non-negativity constraint.

The upshot of all this is that in the context of our analysis it makes sense to impose an upper bound threshold on  $a$ , denoted by  $\bar{a}$  such that if  $a \leq \bar{a}$ ,  $t_i^*(m) \geq t_j^*(m)$  is true. In other words, we assume damages to be significant enough for governments such that signatories choose higher taxes than non-signatories, which seems to be the basic motivation to consider border tax adjustments for environmental reasons. So even under the No-BCA-regime we assume damages to be sufficiently strong such that signatories choose higher taxes than non-signatories,  $t_i^*(m) \geq t_j^*(m)$ . For the No-BCA-regime and  $\gamma = 0$ , this implies for instance  $\bar{a} = c + 2n\delta$ . In the same spirit, also under the BCA-regime, we can derive upper bound thresholds  $\bar{a}$  such that if  $a \leq \bar{a}$ ,  $t_i^*(m) \geq t_j^*(m)$  is always true.

Clearly, when comparing the two regimes, we need to assume that parameter  $a$  satisfies both non-negativity constraints, thus the joint lower bound  $\underline{a}$  is the maximum of the two lower bounds. Also regarding the upper bound  $\bar{a}$ , the joint upper bound is the minimum of the two upper bounds.

Already under the No-BCA-regime, equilibrium taxes are huge terms, which are even bigger under the BCA-regime. Even though we refer the reader to the analytic results for the No-BCA-regime in Finus and Al Khourdajie (2018), we have not been able to do the same for the BCA-regime. Hence, as we want to compare both regimes, we need to resort to simulations, provide all detailed simulation results, conducted with the mathematical software program Maple, upon request and list all qualitative results on which we comment in Section 5 in Appendix 2.

In the simulations, we consider three values of the TFV parameter: no TFV with  $\gamma = 1$ , partial TFV with  $\gamma = 0.5$ , and full TFV with  $\gamma = 0$ . We also consider three values for the damage parameter  $\delta = \{10, 50, 100\}$ , assume  $n = 10$  countries and normalize the cost parameter by setting  $c = 0$ . We recall that we assume  $\phi = 1$  for the adjustment parameter for the reason provided above. For these parameter values, the joint lower and upper bounds of parameter  $a$  are listed in the table below.

	$\delta = 10$		$\delta = 50$		$\delta = 100$	
	$\underline{a}$	$\bar{a}$	$\underline{a}$	$\bar{a}$	$\underline{a}$	$\bar{a}$
$\gamma = 0$	101	104	501	520	1001	1040
$\gamma = 0.5$	120	1000	590	5300	1250	10500
$\gamma = 1$	250	5000	1350	25000	2500	50000

As  $\underline{a}(\gamma, \delta) \leq a(\gamma, \delta) \leq \bar{a}(\gamma, \delta)$ , each parameter space is divided into 6  $a$ -values with equidistant intervals such that  $\Delta(\gamma, \delta) = \frac{\bar{a}(\gamma, \delta) - \underline{a}(\gamma, \delta)}{5}$ ,  $a_1(\gamma, \delta) = \underline{a}(\gamma, \delta)$ ,  $a_2(\gamma, \delta) = \underline{a}(\gamma, \delta) + \Delta(\gamma, \delta)$ , ...,  $a_6(\gamma, \delta) = \underline{a}(\gamma, \delta) + 5\Delta(\gamma, \delta) = \bar{a}(\gamma, \delta)$ . So for instance for  $\gamma = 0$  and  $\delta = 10$ , we have  $a_1(\gamma, \delta) = 101$ ,  $a_2(\gamma, \delta) = 101.6$ , ...,  $a_6(\gamma, \delta) = 104$ . Roughly speaking, for a given  $\gamma$ , moving from  $\delta = 10$  to  $\delta = 50$ , all bounds are inflated by a factor of 5 and by moving to  $\delta = 100$ , they are inflated by a factor of 10. So the different values of the damage parameter  $\delta$  can be considered as a sensitivity analysis for which all qualitative results on which we report below are confirmed to be robust. (So absolute values of the parameters  $a$ ,  $c$  and  $\delta$  do not seem to matter but their ratio.) Subsequently, in this but also the subsequent sections, we are interested in the main qualitative results, not so much in the detailed quantitative results.

In a first step, it is informative to compare equilibrium taxes under both regimes.

## Result 1 - Comparing Equilibrium Taxes Across Regimes

Denote equilibrium taxes under both regimes with superscript No-BCA and BCA, respectively, and assume  $1 < m < n$ .

- Under the BCA-regime signatories' equilibrium taxes are higher than under the No-BCA-regime:  $t_i^{*BCA} > t_i^{*No-BCA} \forall m$ .
- Under the BCA-regime non-signatories' equilibrium taxes are higher than under the No-BCA-regime for  $\gamma = \{0, 0.5\}$ :  $t_j^{*BCA} > t_j^{*No-BCA} \forall m$ . For  $\gamma = 1$ ,  $t_j^{*BCA} < t_j^{*No-BCA} \forall m \leq \tilde{m}$ , and  $t_j^{*BCA} > t_j^{*No-BCA} \forall m > \tilde{m}$ , with  $\tilde{m}$  some threshold,  $\tilde{m} < n$ .
- Under the BCA-regime, total output is lower and hence total emissions are lower than under No-BCA-regime for all  $m$ .

The BCA-policy provides signatory governments with an additional strategic tool to internalize externalities from emissions but also to protect their firms' competitiveness in signatory markets. Furthermore, and importantly, it also serves as an additional source for revenues, the collection of tariffs imposed on imports. Therefore, taxes of signatory governments are higher under the BCA-regime than under the No-BCA-regime.

From non-signatory governments' point of view, BCAs have the following implications. Firstly, their consumers need to pay higher prices for varieties supplied by firms located in signatory countries. Secondly, their firms face an additional tax burden at the borders to signatories' markets that will negatively affect their profits. Thirdly, they face a loss of potential tax revenue, i.e. tax revenue generated by their firms but of which some portion goes into signatory governments' coffers. Therefore, non-signatory governments reaction is complex. On the one hand, non-signatory governments could raise their taxes in order to protect their tax revenues. On the other hand, they could lower their taxes to protect their consumers.

The incentive to protect domestic consumers decreases the lower the value of  $\gamma$  and vanishes for the full TFV with  $\gamma = 0$ . In our simulations, this incentive is also sufficiently low for  $\gamma = 0.5$ , such that for the full and partial TFV, i.e.  $\gamma = \{0, 0.5\}$ , the revenue protection effect dominates the consumer protection effect and also non-signatory governments choose a higher tax under the BCA- than under the No-BCA-regime. For no TFV, i.e.  $\gamma = 1$ , this is also true if  $m > \tilde{m}$ , but is reversed if  $m \leq \tilde{m}$ . However, even in the last case, the overall tax level under the BCA-regime is higher than under the No-BCA-regime such that total output and hence total emissions are always lower for every coalition size (which does not include the grand coalition because then No-BCA- and BCA-regime coincide).

## 5 First Stage

### 5.1 Results Open Membership

In this subsection, we derive stable coalitions under the No-BCA- and BCA-regime as a Nash equilibrium in the simultaneous move open membership single coalition game. We also evaluate those outcomes in terms of total welfare. In subsection 5.2, we provide a rationale for those results by considering various properties of our coalition game. We then consider in subsection 5.3 whether and how our results would change for a sequential coalition formation process and exclusive membership. All results are supported by Appendix 2, which provides more detailed results than the aggregate result, which we present in Results 2 and 3 below.

In order to evaluate the outcome in the two regimes, we consider a relative welfare measure proposed in Eyckmans and Finus (2006) called the closing the gap index (CGI), which is defined as follows:

$$CGI(m^*) := \frac{\sum_{k \in N} Wq_k(m^*) - \sum_{k \in N} Wq_k(m=1)}{\sum_{k \in N} Wq_k(m=n) - \sum_{k \in N} Wq_k(m=1)} \bullet 100. \quad (21)$$

This index measures to which extent a stable agreement with  $m^*$  members closes the gap between the grand coalition, a coalition including all countries ( $m = n$ ), corresponding to the social optimum, and no agreement ( $m = 1$ ), corresponding to the non-cooperative equilibrium, the classical Nash equilibrium without coalition formation. The CGI expresses this in percentage terms. Hence for instance if the grand coalition is a stable agreement, the CGI is 100% whereas if no agreement is stable, i.e. no non-trivial coalition with at least two members is stable, the CGI is 0%.<sup>12</sup>

## Result 2 - Equilibrium Coalitions under Open Membership

*Let  $m^*$  denote the equilibrium size of a stable coalition under open membership and let  $CGI(m^*)$  denote the closing the gap index of stable agreements with  $m^*$  members as defined in equation (21). Then, under the No-BCA- and BCA-regime, we find:*

Table 1 to be inserted here

Result 2 clearly supports two messages. Under the No-BCA-regime stable agreements are small if they exist at all. Only if the taste for variety parameter  $\gamma$  is sufficiently small will an agreement with at least two members be stable. But even for  $\gamma = 0$ , only an agreement with three countries is stable in our model.<sup>13</sup> Accordingly, the closing the gap index is at best small if not zero (see Table 4, Appendix 2 for further details). In our model, the strategic interaction among markets and hence governments is related to the TFV-parameter  $\gamma$ : the larger  $\gamma$ , the lower the taste for variety, the larger the strategic interaction. Under the No-BCA-regime, strong strategic interaction means strong free-rider incentives. In contrast, under the BCA-regime, large stable coalitions are possible, including the grand coalition, with accordingly large values for the closing the gap index. BCAs put signatories in a strong position, which is particularly useful for them if there is strong strategic interaction between signatories and non-signatories. For  $\gamma = 1$  and  $\gamma = 0.5$  we find the grand coalition to be stable in our simulation runs (and hence the CGI is 100%), whereas for  $\gamma = 0$ , depending on the parameter values, we find a coalition ranging from  $m^* = 6$  to  $m^* = 9$  with a CGI still substantially above 85%. See Table 10 in Appendix 2 for further details.

---

<sup>12</sup>In principle, the nominator may be negative and hence also the CGI. In these particular cases, it seems sensible to indicate only that the CGI is negative but not to provide exact numbers. See Result 4 below, in particular Table 2.

<sup>13</sup>For an analytic proof under the No-BCA-regime, see Finus and al Khourdjie (2018) for  $\gamma \in \{0, 1\}$ ; a proof for  $\gamma = 0.5$  is available upon request from the authors.

## 5.2 Properties

In this subsection, we want to rationalize Result 2 above. For this we first define some properties and then report to which extent these properties hold in the two regimes. The notation makes use of the fact that welfare depends only on membership and on the size of coalitions  $m$ , but not on the composition of individual members due to our assumption of ex-ante symmetric players, i.e. all players have the same welfare function, and hence all signatories and all non-signatories have the same welfare, though signatories and non-signatories have different welfare because they choose different equilibrium strategies.

Let  $1 \leq m \leq n - 1$ :

- **Positive Externality:** the move from  $m$  to  $m + 1$  exhibits a (strict) positive externality if:

$$W_{j \notin S}(m + 1) \geq (>) W_{j \notin S}(m)$$

- **Negative Externality:** the move from  $m$  to  $m + 1$  exhibits a strict negative externality if:

$$W_{j \notin S}(m + 1) < W_{j \notin S}(m)$$

- **Superadditivity:** the move from  $m$  to  $m + 1$  is (strictly) superadditive if:

$$[m + 1] W_{i \in S}(m) \geq (>) m W_{i \in S}(m) + W_{j \notin S}(m)$$

- **Global Cohesiveness:** a coalition game is globally (strictly) cohesive if:

$$n W(n) \geq (>) m W(m) + [n - m] W_j(m) \forall m \neq n$$

- **Cohesiveness:** the move from  $m$  to  $m + 1$  is (strictly) cohesive if:

$$[m + 1] W_{i \in S}(m + 1) + [n - m - 1] W_{j \notin S}(m + 1) \geq (>) m W_{i \in S}(m) + [n - m] W_{j \notin S}(m)$$

The first four properties are related to positive features of coalition formation, the last two properties are related to the normative dimension of coalition formation.<sup>14</sup>

In positive externality games, outsiders benefit from the enlargement of the agreement. This makes it attractive to remain a non-signatory. It is for this reason why in many positive externality games, stable coalitions are small. This may be the

---

<sup>14</sup>For a discussion of these properties and other economic examples, see Bloch (2003) and Yi (1997, 2003).



case despite superadditivity holds, which provides an incentive to join a coalition. Simply speaking, in many positive externality games, for larger coalitions, the superadditivity effect is smaller than the positive externality effect. However, there are games, in which even superadditivity fails to hold, in particular if coalition  $S$  is small. Public provision games are typical examples. By enlarging the coalition, signatories reduce their emissions from which also non-signatories benefit (positive externality game). Moreover, if emissions are strategic substitutes, non-signatories increase their emissions as an optimal response. If this leakage is strong enough, superadditivity may fail, which is particular the case for small coalitions, as there are many non-signatories which free-ride. The formation of output cartels exhibits similar features. Firms in the cartel reduce their output in order to stabilize prices from which also firms outside the cartel benefit. Their optimal response calls for an increase in output, which may cause superadditivity to fail. Our strategic trade-environment game combines the features of an output cartel and a public good game and hence exhibits similar features under the No-BCA-regime. In contrast, in negative externality games, outsiders have an incentive to join coalition  $S$ . It is easy to show that in a game which is fully superadditive (i.e. superadditivity holds for every move  $m$  to  $m + 1$ ) and exhibits a full negative externality (i.e. every move  $m$  to  $m + 1$  is associated with negative externalities), the grand coalition is the unique stable coalition in the simultaneous move open membership single coalition game (Weikard 2009). Under the BCA-regime, there are cases where the coalition game exhibits negative externalities, which we discuss in detail below.

Global cohesiveness simply means that global welfare in the grand coalition is larger than in any other coalition. This property holds strictly in any externality game and hence also in our strategic trade-environment game as all externalities are internalized in the grand coalition by assumption. A more specific property is cohesiveness which implies that the move from a coalition of size  $m$  to  $m + 1$  is associated with a global welfare gain. If this holds for every move, we say the game is fully cohesive, i.e. global welfare increases continuously each time a coalition is enlarged. Clearly, in a game which such a feature, there is a normative motivation to search for large stable coalitions even if the grand coalition is not stable. Note that in games, which exhibit a full positive externality and which are fully superadditive, full cohesiveness is implied: every enlargement benefits outsiders and those players involved in the enlargement gain at the aggregate. The No-BCA-regime generally displays this property. In contrast, in negative externality games, it is less obvious that full cohesiveness holds. Though the grand coalitions generates the largest global welfare, global welfare may not continuously increase with the enlargement of coalitions and in fact there may even be intermediate coalitions with lower welfare than in the status quo with no agreement. This may happen if the negative externality effect is very strong. As we will discuss below, the BCA-regime may show this feature.

Equipped with these definitions, we now analyze the driving forces under the two regimes which explains Result 2. Appendix 2 supports the discussion with detailed

results.

Under the No-BCA-regime, we find that the full positive externality property holds generally. Whenever the coalition is enlarged, signatories decrease their outputs whereas non-signatories expand their outputs (as long as  $\gamma > 1$ ), though the overall output (and hence environmental damages) decrease. So non-signatories benefit from lower damages, higher prices and a relocation of production, compensating for a possible loss of consumer surplus due to high taxes by signatory governments. Full superadditivity holds for  $\gamma = 0.5$  and  $\gamma = 0$  but fails for  $\gamma = 1$  (see Table 1, Appendix 2). In fact, for  $\gamma = 1$ , for every move from  $m$  to  $m + 1$  superadditivity fails, except for the last move from  $n - 1$  to  $n$ . The reason is that the leakage effect increases with the value of  $\gamma$ . Given that superadditivity is a necessary condition for internal stability in positive externality games, it is not surprising that we found that no non-trivial coalition is stable for  $\gamma = 1$ . But even if full superadditivity holds, as for  $\gamma = 0.5$  and  $\gamma = 0$ , positive externalities are even stronger, which explains small stable coalitions with  $m^* = 2$  and  $m^* = 3$ , respectively. Also note that in line with other positive externality games with ex-ante symmetric players, non-signatories are always better off than signatories,  $W_{i \in S}(m) < W_{j \notin S}(m)$  for every  $m$ ,  $1 < m < n$ . All countries enjoy the same consumer surplus and suffer the same damages but gross profits are lower in signatory countries due to the lower outputs of their firms. Hence, the coalition game resembles a kind of n-player (symmetric) chicken game (in pure strategies) with either no or a small number of chickens in equilibrium. Given that that full cohesiveness holds generally in the No-BCA-regime, stresses the well-known paradox of cooperation mentioned in the Introduction. Despite larger coalitions would be associated with a global welfare gain even if the grand coalition cannot be obtained, only small coalitions are stable which cannot or hardly can improve upon the status quo without agreement.

Under the BCA-regime, we find that full superadditivity holds throughout all simulations as signatories can now better control the leakage effect. For  $\gamma = 0$ , we also have full positive externalities. However, for  $\gamma = 0.5$  and  $\gamma = 1$ , we do not have generally positive externalities, in fact, only if the damage parameter  $\delta$  is very high compared to the demand parameter  $a$ , i.e. for very small values of  $a(\gamma, \delta)$ , and small coalitions, i.e. small  $m$  (see Table 6, Appendix 2). That is, non-signatories enjoy positive externalities because the reduction of damages matters a lot and the strategic advantage of signatories is still moderate due to smaller agreements and/or small  $\gamma$ . However, if the reduction of damages is slightly less important, coalitions are sufficiently large and/or  $\gamma$  is sufficiently large so that the strategic advantage of signatories is sufficiently large, then signatories impose negative externalities on outsiders through BCAs. These are situations in which signatories extract high tariff revenues from non-signatories, consumers in non-signatory countries are hit hard due to high taxes of signatory governments but non-signatories do not benefit sufficiently from lower damages to compensate these negative externalities. This is one line of arguments which explains why large stable coalitions emerge under the BCA-regime. Another,

and additional line of arguments emerges by noting that  $W_{i \in S}(m) < W_{j \notin S}(m)$  holds now only for some larger coalitions for  $\gamma = 0$ , but is reversed in all other cases, i.e.  $W_{i \in S}(m) > W_{j \notin S}(m)$  (see Table 5, Appendix 2). It is also interesting to note that signatories are always better off under the BCA- than under No-BCA-regime ( $W_{i \in S}^{BCA}(m) > W_{i \in S}^{N0-BCA}(m)$  for all  $m$ ,  $1 < m < n$ ), and that for non-signatories this is reversed ( $W_{j \notin S}^{BCA}(m) < W_{j \notin S}^{N0-BCA}(m)$  generally for  $\gamma = 1$  and  $\gamma = 0.5$  and for  $\gamma = 0$  at least for large coalitions (see Table 11, Appendix 2). So compared to the No-BCA-regime, with border carbon adjustments the situation of signatories improves and that of non-signatories worsens, leading to larger stable agreements. In terms of global welfare, for  $\gamma = 1$  and  $\gamma = 0.5$  the grand coalition is stable and hence the optimum is achieved due to the (general) property global cohesiveness and for  $\gamma = 0$  full cohesiveness holds (see Table 8, Appendix 2) and hence any coalition larger than  $m^* = 3$  as under the No-BCA-regime is associated with larger global welfare under the BCA-regime with  $m^*$  ranging from 6 to 9 (i.e.  $m^* = 6 - 9$  in Table 1).

### 5.3 Alternative Coalition Formation Games

In this subsection, we want to analyze how robust our conclusions are regarding the assumptions of our coalition formation game. Apart from this technical point, and even more important, we are interested in what would change if agreements are not open to accession of outsiders (open membership) but are subject to an approval process by current members (exclusive membership). Therefore, we consider two alternative features: a) a sequential instead of a simultaneous coalition formation process and b) exclusive instead of open membership. We may note that though environmental agreements are typical of the open membership type, trade agreements classically feature exclusive membership, including accession to the WTO or the European market.

The standard assumption in the literature on coalition formation is a simultaneous coalition formation (see Bloch 2003 and Yi 1997, 2003). However, in reality, agreements typically form sequentially, with some initiators moving first and some laggards joining later. The simplest extension of our game could assume that two countries form a coalition in a first instance. In a sequential process others may join if they benefit from accession. In this sequential process, any intermediate coalition of size  $m$  must first of all be internally stable. Outsiders will accede to an intermediate coalition of size  $m$  if this coalition is not externally stable. However, in the case of symmetric players, this implies that a coalition of size  $m + 1$  is strictly internally stable. Thus, if there is a sequence of coalitions  $m$  in the interval  $1 < m \leq \bar{m} \leq n$  which are all strictly internally stable ( $W_{i \in S}(m) - W_{i \notin S}(m - 1) > 0$ ), then there will be a sequence of accessions until  $\bar{m}$  is reached. If  $\bar{m} = n$ , then  $\bar{m}$  is externally stable by definition (as no outsider is left) and if  $\bar{m} < n$ , and  $m = \bar{m} + 1$  is internally unstable, then  $\bar{m}$  is also externally stable. Hence, in such a sequence,  $m^* = \bar{m}$  emerges.

Technically, we can summarize internal stability,  $W_{i \in S}(m) - W_{i \notin S}(m-1) \geq 0$ , and external stability,  $W_{j \notin S}(m) - W_{j \notin S}(m+1) \geq 0$  (under open membership) into a stability function  $\Phi(m) := W_{i \in S}(m) - W_{i \notin S}(m-1)$  with a coalition of size  $m$  being internally stable if  $\Phi(m) \geq 0$  and externally stable if  $\Phi(m+1) < 0$ . Tables 3 and 9 in Appendix 2 display the result for the stability function for the No-BCA- and BCA-regime, respectively. For both regimes,  $\Phi(m) > 0$  for a sequence of coalitions,  $m$ ,  $1 < m \leq \bar{m} \leq n$ , implying that all coalitions smaller than  $\bar{m}$  are externally unstable except  $\bar{m}$  and hence  $m^* = \bar{m}$ . (For  $\gamma = 1$  under the No-BCA-regime, we have  $\Phi(m) < 0$  for all  $m > 1$  and hence  $m^* = 1$  is obvious.) The upshot of all this is that the results obtained above and summarized in Result 2 for a simultaneous coalition formation process would be exactly the same for a sequential coalition formation process and hence our results are robust regarding this modification of assumption.

Let us now move to the second alternative assumption of exclusive membership. Consider first our original assumption of a simultaneous coalition formation process. Of course, also for exclusive membership, a coalition of size  $m$  can only be stable if it is internally stable, i.e.  $\Phi(m) \geq 0$ . Moreover, if the enlarged coalition  $m+1$  is not internally stable, then  $m$  is externally stable. That is, outsiders have no incentive to join coalition  $m$  anyway and hence like under open membership coalition  $m$  is externally stable. However, if to the contrary,  $m$  is externally unstable under open membership (and hence  $m+1$  is strictly internally stable), exclusive membership could make a difference. The reason is that the current members of coalition of size  $m$  would only agree to accession if and only if they are better off. Signatories are better off (worse off) if they enjoy a positive internal spillover (negative internal spillover).

Let  $1 \leq m \leq n-1$ :

- **Positive Internal Spillover:** the move from  $m$  to  $m+1$  exhibits (strict) positive internal spillovers if:

$$W_{i \in S}(m+1) \geq (>) W_{i \in S}(m)$$

- **Negative Internal Spillover:** the move from  $m$  to  $m+1$  exhibits strict negative internal spillovers if:

$$W_{i \in S}(m+1) < W_{i \in S}(m)$$

There are two main reasons for possible negative internal spillovers in our model. Under the No-BCA-regime, though more countries joining an agreement means less free-riders, it also means more ambitious targets, which in the context of strong leakage effects may lower the welfare of signatories. We observe negative internal spillovers under the No-BCA-regime for small coalitions and  $\gamma = 1$  (see Table 2, Appendix 2). Under the BCA-regime there is another, very “selfish” reason. Signatories enjoy a strong position towards non-signatories. If the number of signatories

has reached some level  $\tilde{m}$ , the basis to extract tariff revenues from non-signatories becomes smaller, so that the current members do not benefit from a further expansion of the agreement. We find negative internal spillovers under the BCA-regime for  $\gamma = 1$  and  $\gamma = 0.5$  for a threshold coalition size  $\tilde{m}$  which can be substantially smaller than the grand coalition (see Table 7, Appendix 2).

Now, under the No-BCA-regime, exclusive membership will not make a difference because  $\tilde{m}$  is much larger than the largest internally stable coalition  $\bar{m}$ . For  $\gamma = 1$ ,  $\Phi(m) < 0$  for all  $m > 1$  and hence all larger coalitions are anyway externally stable and so  $m^* = 1$ . For  $\gamma = 0.5$ ,  $\tilde{m} = 10$ ,  $\Phi(m) > 0$  for  $m = 2$  but  $\Phi(m) < 0$  for all  $m > 2$  and hence  $m^* = 2$ . For  $\gamma = 0$ , again,  $\tilde{m} = 10$ ,  $\Phi(m) > 0$  for  $m = 2$  and  $m = 3$  but  $\Phi(m) < 0$  for all  $m > 3$  and hence  $m = 2$  is not externally stable whereas  $m = 3$  is so that  $m^* = 3$  (see Table 3, Appendix 2). Hence,  $m^* = \bar{m}$  under exclusive membership as displayed in Result 3 below which is the same as under open membership as reported in Result 2 above.

### Result 3 - Equilibrium Coalitions under Exclusive Membership

*Assume exclusive membership. Let  $m^*$  denote the size of stable coalition(s) under both regimes and consider a sequential coalition formation process. Then, under the No-BCA- and BCA-regime, we find:*

Table 2 to be inserted here

Under the BCA-regime, a more interesting but also more alarming pattern emerges, which is illustrated with an example in Figure 1, which assumes  $\gamma = 1$ ,  $\delta = 10$  and chooses an intermediate value for parameter  $a$ , namely  $a_4(\delta, \gamma) = 3100$ .

Figure 1 about here

For  $\gamma = 1$  and  $\gamma = 0.5$  we recall that all coalitions  $m$  in the interval  $1 < m \leq \bar{m} = n$  (with  $n = 10$  in our simulation runs) are strictly internally stable (see Table 9 in Appendix 2). Consequently, in this interval, non-signatories have an incentive to join any intermediate coalition. For the example, the stability function  $\Phi(m)$  is shown in Figure 1a. Moreover, for smaller coalitions we find positive internal spillovers up to coalitions of size  $\tilde{m} < \bar{m} = n$  and from  $\tilde{m} + 1$  to  $\bar{m} = n$  negative internal spillovers. For the example, Figure 1b illustrates this which plots welfare of signatories as a function of the size of the coalition,  $m$ , with  $\tilde{m} = 5$ . Hence, all coalitions up to  $\tilde{m}$  are not externally stable and all coalitions  $\tilde{m} + 1$  to  $\bar{m} = n$  are externally stable under exclusive membership. Hence for a simultaneous coalition formation process, any coalition  $m$ ,  $\tilde{m} \leq m \leq \bar{m} = n$  is stable and therefore  $m^*$  is not unique for  $\gamma = 1$  and  $\gamma = 0.5$ . In other words, compared to open membership, also smaller coalitions are stable under exclusive membership. In the example,  $m^* = 5 - 10$  (i.e all coalitions of

size 5 to 10 are stable,  $m_L^* = 5$  and  $m_H^* = 10$  with reference to Table 2). In contrast, for  $\gamma = 0$ ,  $\tilde{m} > \bar{m}$  and in fact  $\tilde{m} = n$ , so we have positive internal spillovers for every move up to the grand coalition. So all coalitions up to  $m = \bar{m} - 1$  are externally unstable,  $m = \bar{m}$  is externally stable (because  $\bar{m} + 1$  is internally unstable) and consequently, we have a unique stable coalition  $m^* = \bar{m}$  which is the same as under open membership.

Finally, we may consider what changes if we assume a sequential coalition formation process for exclusive membership. Under the No-BCA-regime it is immediately clear that this will make no difference. The same is true under the BCA-regime and  $\gamma = 0$ , but for  $\gamma = 1$  and  $\gamma = 0.5$  the process of enlargement would stop at  $m = \tilde{m}$  and hence  $m^* = \tilde{m}$  would be the unique stable equilibrium, the smallest stable coalition which we found for a simultaneous coalition formation process, denoted by  $m_L^*$  in Table 2. Thus, a unique equilibrium emerges for a sequential coalition process. In our example,  $m^* = 5$ .

So viewed together, exclusive membership makes no difference for the No-BCA-regime but may lead to smaller stable coalitions for the BCA-regime. This is the case for  $\gamma = 1$  and  $\gamma = 0.5$  where the grand coalition emerges under open membership, but now with exclusive membership, either additional smaller coalitions are stable for a simultaneous coalition formation process or a unique smaller coalition is stable for a sequential coalition formation process.

Important is now to understand how this may affect global welfare. For this we display in Result 3 above for the BCA-regime the CGI for the smallest, denoted  $m_L^*$ , and the largest stable coalition, denoted  $m_H^*$ , for the simultaneous coalition formation process, noting that  $m_L^*$  would be the unique stable coalition size in a sequential coalition formation process. This leads to far less positive conclusions under the BCA-regime. For instance, in our example,  $m^* = \tilde{m} = 5$  in a sequential coalition formation process. However, as is evident from Figure 1c, already at  $m = 3$ , global welfare starts to decline continuously with an enlargement of the coalition until reaching the lowest global welfare at  $m = 8$ , then picking up again until the grand coalition is reached. In this example, global welfare at  $m^* = \tilde{m} = 5$  is even slightly smaller than in the status quo with no agreement. Hence, the CGI would be negative, and the same is true in this example for  $m = 6$  to  $m = 9$  (with even lower global welfare) which would be additionally stable for a simultaneous coalition formation process. Thus, exclusive membership may lead to much worse outcomes than open membership under the BCA-regime, with outcomes which are even worse than if no agreement was reached. As Result 3 suggests for  $\gamma = 1$  and  $\gamma = 0.5$  this is the case for the larger values of parameter  $a$ , namely  $a_4(\delta, \gamma)$  to  $a_6(\delta, \gamma)$ .

The fact that global welfare may decrease in  $m$  for some intermediate coalition sizes means technically that under the BCA-regime not each enlargement is associated with cohesiveness, very different from the No-BCA-regime which is always fully cohesive (see Appendix 2, in particular the summary of the general result under the No-BCA-regime and Table 8 for the BCA-regime). The reason is that under

the BCA-regime the negative externality effect on non-signatories can be so strong that, despite superadditivity, global welfare declines for some intermediate coalitions, which is the case in our simulations for  $\gamma = 1$  and  $\gamma = 0.5$ .

Overall, this allows us to draw at least two important policy conclusions. Firstly, the general fear by most economists that any restriction on trade may cause large welfare losses, even if motivated by environmental concerns, is strongly supported by our model. BCAs may be associated with a global welfare loss if full cooperation is not achieved. Secondly, and nevertheless, border carbon adjustments can be a useful tool to enforce environmental agreements and to increase global welfare but membership should not be restricted because some governments hijack this instrument for their own interest. Border carbon adjustments should support an international environmental agreement and hence should remain open to all countries which would like to join for the global good.

## 6 Summary and Conclusion

In a stylized intra-industry trade model with horizontal product differentiation and taste for variety (TFV) by consumers we studied the formation, stability and success of international climate agreements. In the tradition of the game-theoretic literature on international environmental agreements (IEAs), we modeled a three stage game in which governments first decide whether to join a climate agreement, then decide on their policy levels and finally firms choose their outputs. The model captures the strategic interaction between signatories and non-signatories with an endogenous choice of strategies at each stage. We considered two regimes. Under the first regime, governments have an emission tax at their avail to correct externalities, under the second regime, signatories to a climate agreement have an additional policy option, namely to impose tariffs on imports. The tariffs, the border carbon adjustments (BCAs), are chosen such that effective tax on imports is the same as the tax on domestic production. We labeled the first regime No-BCA-regime and the second BCA-regime. Output-based rebates (OBRs) were not part of the BCA-regime as they are not in accordance with the strong polluters principle, may be difficult to justify within WTO-rules but also because we could show that our BCA-regime is already able to deliver what most scholars intuitively suspect, namely, it enforces larger stable and effective climate agreements.

We showed that BCAs are a game changer in several respects. Firms located in signatory countries, facing higher taxes than their rivals in non-signatory countries, now play on equal terms with their rivals, at least in their home market. Thus, the difference in profits is reduced through BCAs, though it does not vanish. Signatory governments also benefit from tariff revenues, which at the same time constitute a loss of tax revenues to non-signatory governments. However, not all effects work to the absolute or comparative advantage of signatory countries. Without BCAs,

though the consumer surplus is negatively affected by absolute tax levels, consumers in signatory and non-signatory enjoy the same utility. Under the BCA-regime, this changes, consumers in signatories are disadvantaged as they face higher prices than consumers in non-signatory countries. Nevertheless, at the aggregate welfare level of countries, the first two effects dominate the third effect. That is, signatory's welfare increases compared to the No-BCA regime, and in most cases the reverse is true for non-signatory countries.

A variation of this theme showed up in the properties of the coalition formation game. Under the No-BCA regime, non-signatories' welfare increases with the expansion of the agreement, which we called positive externalities. This is a result of the non-excludability of public good provision. Larger climate agreements implies more ambitious emission reduction targets by signatories from which non-signatories benefit at no cost. Moreover, in the context of trade, non-signatories benefit from a relocation of production to their countries. This leakage effect also lowers the gains from cooperation to signatories, reflected in the properties superadditivity and positive internal spillovers. That is, an expansion of an agreement may neither be beneficial for signatories at the aggregate (superadditivity fails) nor individual level (internal spillovers are not positive but negative) if leakage effects are too strong. We showed that this may be particularly true under two conditions. Firstly, as long as participation in an agreement is below a threshold, the number of free-riders is too large. Secondly, if the taste for variety of consumers is low, competition among firms for market shares is fierce, and hence the strategic interaction among governments is particularly strong under the No-BCA-regime. Altogether this explain why under the No-BCA-regime for a sufficiently low taste of variety no climate agreement is stable, but even for higher levels of the taste for variety at best only small agreements are stable. That is, under the No-BCA-regime, agreements can at best only marginally close the gap between no and full cooperation in terms of global welfare.

In contrast, under the BCA-regime, large agreements can be stable, including an agreement comprising all countries (grand coalition). Particularly for those situations with strong strategic interaction among countries, i.e. low taste for variety, the grand coalition is stable. But even for a high taste for variety, stable agreements are much larger than under the No-BCA-regime, and close the gap between no and full cooperation by a substantial amount. BCAs help not only to reduce carbon leakage, but extract tax revenues from firms located in non-signatory countries. The game change implied that superadditivity always holds and that non-signatories may no longer enjoy positive externalities but may suffer from negative externalities if agreements are sequentially increased. Thus, BCAs strengthen the incentive of countries to join a climate agreement. It is for this reason that irrespective of whether coalition formation is modeled as a one-shot or sequential game, self-enforcing climate agreements are large and successful (in welfare terms) under open membership. That is, if accession to an agreement cannot be denied, as this is common practice for almost all international environmental agreements, BCAs accomplish for what they



are designed. That is, in equilibrium, either BCAs only serve as a threat but are not implemented under full participation, or they are only imposed on a small group of outsiders. However, we also showed that for smaller agreements of intermediate size, the welfare cost of BCAs may be very large, sometimes implying lower global welfare than in the status quo with no agreement. This may happen if negative externalities to non-signatories are higher than the benefits to signatories, which is particularly likely in our model if the strategic interaction among countries is strong, so low levels for the taste for variety. We showed that if climate agreements, in the tradition of most trade agreements, are of the exclusive membership type (i.e. current signatories vote on the application for accession by outsiders), smaller stable agreements may emerge, with possible negative impacts for global welfare as described above. The reason is that despite outsiders may want to join the climate agreement, this may not be in the strategic interest of current signatories. From a pure selfish point of view, signatories' welfare may only increase up to some threshold membership and then declines. Above the threshold, any additional reduction of environmental damages and further effort to strengthen domestic firms' competitiveness is not outweighed by the reduction of tariff revenues, as result of a shrinking tariff revenue base. Thus, in order to avoid that BCAs serve the narrow interest of a few countries, it clearly emerges from our analysis that climate agreements should remain open to all countries.

Our analysis made some simplifying assumptions. In the following, we briefly discuss the implication of relaxing these assumptions. Firstly, we did not consider OBRs for the reasons given above. With OBRs, signatories would gain an even strong position towards non-signatories, which, as we have shown, is not necessary to make BCAs successful. We expect that the possible negative impacts under exclusive membership would be more pronounced. Secondly, we could allow for asymmetric welfare functions of countries. Though more realistic, the driving forces/properties identified in our model with ex-ante symmetric players would not disappear. If tariff and tax revenues were used to compensate those benefiting from cooperation less than others, asymmetry does not have to be an obstacle but may be an asset for cooperation as this emerges from asymmetric IEA models for instance by Finus and McGinty (2015) and Weikard (2009). Thirdly, our model allowed for the relocation of production but did not consider the relocation firms as a result of different taxes in signatory and non-signatory countries. Hence, an interesting extension could build on strategic models with endogenous plant location, like in Hoel (1997), Markusen et al. (1993) and Motta and Thisse (1994) by adding the perspective of BCAs. We expect that without transportation and relocation costs, the knife-edge equilibrium in these models with a race to the bottom would not disappear with border adjustments only and hence BCAs would be only successful if complemented by other trade measures like OBRs.

## Acknowledgments

We would like to acknowledge the helpful comments by Sergio Currarini, Lucy O'Shea, Javier Rivas, Santiago Rubio and Aart de Zeeuw. Alaa al Khourdajie would like to thank in particular Pedro Pintassilgo for his hospitality during a research visit at the University of the Algarve, Faro, Portugal in 2016 and the many fruitful discussions. The authors bear the responsibility of any errors and omissions that may remain.

## 7 References

1. Anouliès L. (2014), The Strategic and Effective Dimensions of Border Tax Adjustment. "Journal of Public Economic Theory", vol. 17(6), pp. 824-847.
2. d'Aspremont, C., A. Jacquemin, J.J. Gabszewicz and J.A. Weymark (1983), On the Stability of Collusive Price Leadership. "Canadian Journal of Economics", vol. 16(1), pp. 17-25.
3. Baksi, S. and A.R. Chaudhuri (2014), Trade, Border Tax Adjustments, and the Stability of International Environmental Agreements. Working Paper.
4. Baksi, S. and A.R. Chaudhuri (2017), International Trade and Environmental Cooperation among Heterogeneous Countries. In: O. Kayalca, S. Cagatay and H. Mihci (eds.), Economics of International Environmental Agreements: A Critical Approach, Routledge, pp. 97-115.
5. Barrett, S. (1994a), Self-Enforcing International Environmental Agreements. "Oxford Economic Papers", vol. 46, pp. 878-894.
6. Barrett, S. (1994b), Strategic Environmental Policy and International Trade. "Journal of Public Economics", vol. 54, pp. 325-338.
7. Bloch, F. (2003), Non-Cooperative Models of Coalition Formation in Games with Spillovers. In: Carraro, C. (ed.), The Endogenous Formation of Economic Coalitions. Edward Elgar, Cheltenham, UK et al., ch. 2, pp. 35-79.
8. Böhringer, C., E.J. Balistreri, and T.F. Rutherford (2012), The Role of Border Carbon Adjustment in Unilateral Climate Policy: Overview of an Energy Modeling Forum study (EMF 29). "Energy Economics", vol. 34, pp. 97-110.
9. Böhringer, C., B. Bye, T. Fæhn, and K. E. Rosendahl (2017a), Output-based Rebating of Carbon Taxes in the Neighbors Backyard. "Canadian Journal of Economics", vol. 50(2), pp. 426-455.

10. Böhringer, C. A. Lange, and T. F. Rutherford (2014), Optimal Emission Pricing in the Presence of International Spillovers: Decomposing Leakage and Terms-of-trade Motives. “Journal of Public Economics”, vol. 110, pp. 101-111.
11. Böhringer, C., A. Müller, and J. Schneider (2015), Carbon Tariffs Revisited. “Journal of the Association of Environmental and Resource Economists”, vol. 2(4), pp. 629-672, 2015.
12. Böhringer, C. K. E. Rosendahl and H. B. Storrøstenc (2017b), Robust Policies to Mitigate Carbon Leakage. “Journal of Public Economics”, vol. 149, pp. 35-46.
13. Brander, J.A. and B.J. Spencer (1985), Export Subsidies and International Market Share Rivalry. “Journal of International Economics”, vol. 18, pp. 83-100.
14. Branger, F. and P. Quirion (2014), Would Border Carbon Adjustments Prevent Carbon Leakage and Heavy Industry Competitiveness Losses? Insights from a Meta-analysis of Recent Economic Studies. “Ecological Economics”, vol. 99(C), pp. 29-39.
15. Caron, J. (2012), Estimating Carbon Leakage and the Efficiency of Border Adjustments in General Equilibrium — Does Sectoral Aggregation Matter? “Energy Economics”, vol. 34, pp. 111-126.
16. Carraro, C. and D. Siniscalco (1993), Strategies for the International Protection of the Environment. “Journal of Public Economics”, vol. 52(3), pp. 309-328.
17. Conrad, K. (1993). Taxes and Subsidies for Pollution-Intensive Industries as Trade Policy. “Journal of Environmental Economics and Management” 25, pp. 121-135.
18. Copeland, B. R. and M. S. Taylor (1995), Trade and Transboundary Pollution. “American Economic Review”, vol. 84(4), pp. 851-874.
19. Di Comite, F., J-F. Thisse and H. Vandenbussche (2014), Verti-zontal Differentiation in Export Markets. “Journal of International Economics”, vol. 93(1), pp. 50-66.
20. Dissou, Y. and T. Eyland (2011), Carbon Control Policies, Competitiveness, and Border Tax Adjustments. “Energy Economics”, vol. 33, pp. 556-564.
21. Dixit, A. K. and J. Stiglitz (1977), Monopolistic Competition and Optimum Product Diversity. “The American Economic Review”, vol. 67(3), pp. 297-308.
22. Eichner, T. and R. Pethig (2013), Self-Enforcing Environmental Agreements and International Trade. “Journal of Public Economics”, vol. 102, pp. 37-50.

23. Eichner, T. and R. Pethig (2014), Forging a Global Environmental Agreements Through Trade Sanctions on Free Riders? Discussion Papers in Economics, University of Siegen, Faculty of Economics, No. 171-14.
24. Eichner, T. and R. Pethig (2015), Self-Enforcing International Environmental Agreements and Trade: Taxes Versus Caps. Oxford Economic Papers vol. 67, pp. 897-917.
25. Eyckmans, J. and M. Finus (2006), New Roads to International Environmental Agreements: The Case of Global Warming. "Environmental Economics and Policy Studies", vol. 7, pp. 391-414.
26. Eyland, T. and G. Zaccour (2013), Strategic Effects of a Border Tax Adjustment. "International Game Theory Review", vol. 14(03).
27. Eyland, T. and G. Zaccour (2014), Carbon Tariffs and Cooperative Outcomes. "Energy Policy", vol. 65, pp. 718-728.
28. Financial Times (Feb. 12, 2017), A Carbon Tax is the Best Answer to Climate Change. Lakshmi Mittal.
29. Finus, M. and A. Al Khourdajie (2017), Strategic Environmental Policy, International Trade and Self-enforcing Agreements: The Role of Consumers' Taste for Variety. Bath Economic Research Papers, No. 60/17. Forthcoming "Strategic Behavior and the Environment".
30. Finus, M. and A. Caparrós (2015), Game Theory and International Environmental Cooperation: Essential Readings. Edward Elgar, Cheltenham.
31. Finus, M. and M. McGinty (2015), The Anti-Paradox of Cooperation: Diversity Pays. Bath Discussion Paper Series, 40/2015.
32. Fischer C. (2015), Options for Avoiding Carbon Leakage. In: Barrett, S., C. Carraro and J. de Melo (eds.), Towards a Workable and Effective Climate Regime. Chapter 21, pp. 297-314, CEPR Press, London.
33. Fischer C. and A.K. Fox (2012), Comparing Policies to Combat Emissions Leakage: Border Carbon Adjustments versus Rebates. "Journal of Environmental Economics and Management", vol. 64, pp.199–216.
34. Helm, D., C. Hepburn and G. Ruta (2012), Trade, Climate Change, and the Political Game Theory of Border Carbon Adjustments. "Oxford Review of Economic Policy", vol. 2, pp. 368-394.
35. Helpman, E. and P.R. Krugman (1985), Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy. The MIT Press. Cambridge, Massachusetts. London, UK.

36. Hoel, M. (1996), Should a Carbon Tax be Differentiated Across Sectors? “Journal of Public Economics”, vol. 59, pp. 17-32.
37. Hoel, M. (1997), Environmental Policy with Endogenous Plant Location. “Scandinavian Journal of Economics”, vol. 99, pp. 241-259.
38. Irfanoglu, Z.B., J.P. Sesmero, and A. Golub (2015), Potential of Border Tax Adjustments to Deter Free Riding in International Climate Agreements. “Environmental Research Letters”, vol. 10(2).
39. Jakob, M., R. Marschinski and M. Hübler (2013), Between a Rock and a Hard Place: A Trade-Theory Analysis of Leakage under Production and Consumption-Based Policies. “Environmental and Resource Economics”, vol. 56, pp. 47-72.
40. Keen, M. and C. Kotsogiannis (2014), Coordinating Climate and Trade Policies: Pareto Efficiency and the Role of Border Tax Adjustments. “Journal of International Economics”, vol. 94(1), pp. 119-128.
41. Kennedy, P.W. (1994), Equilibrium Pollution Taxes in Open Economies with Imperfect Competition. “Journal of Environmental Economics and Management”, vol. 27, pp. 49-63.
42. Loke, W-H. and L.A. Winters (2012), Why is Open Regionalism Not Always Good? “Journal of Economic Integration”, vol. 27(3), pp. 437-465.
43. Markusen, J.R. (1975), International Externalities and Optimal Tax Structures. “Journal of International Economics”, vol. 5, pp. 15-29.
44. Markusen, J. R., E. R. Morey and N. Olewiler (1993), “Environmental Policy when Market Structure and Plant Locations Are Endogenous. “Journal of Environmental Economics and Management”, vol. 24, pp. 69-86.
45. Mehling, M., H. van Asselt, K. Das, S. Dröge and C. Verkuijl (2017), Designing Border Carbon Adjustments for Enhanced Climate Action, Climate Strategies.
46. Motta, M. and J.-F. Thisse (1994), Does Environmental Dumping Lead to Delocation? “European Economic Review”, vol. 38, pp. 563-576.
47. Stiglitz, J.E. (2006), A New Agenda for Global Warming. Economists’ Voice. July 2016,, 3(7). [https://www.carbontax.org/wp-content/uploads/2007/09/stiglitz\\_-\\_a-new-agenda-for-global-warming\\_-\\_economists-voice\\_-\\_july-2006.pdf](https://www.carbontax.org/wp-content/uploads/2007/09/stiglitz_-_a-new-agenda-for-global-warming_-_economists-voice_-_july-2006.pdf)
48. Tsakiri, N., M. S. Michael and P. Hatzipanayotou (2014), Asymmetric Tax Policy Responses in Large Economies With Cross-Border Pollution. “Environmental and Resource Economics”, vol. 58, pp. 563-578.

49. Ulph, A. (1996), Environmental Policy and International Trade when Governments and Producers Act Strategically. “Journal of Environmental Economics and Management”, vol. 30, pp. 265-281.
50. Vlassis, N. (2013), The Welfare Consequences of Pollution-Tax Harmonization. “Environmental and Resource Economics”, vol. 56, pp. 227-238.
51. Weikard, H.-P. (2009), Cartel Stability under an Optimal Sharing Rule. “Manchester School”, vol. 77(5), pp. 575-593.
52. Weitzel, M., M. Hübner and S. Peterson (2012), Fair, Optimal or Detrimental? Environmental vs. Strategic Use of Border Carbon Adjustment. “Energy Economics”, vol. 34, pp. 198-207.
53. Yi, S.-S. (1996), Endogenous Formation of Customs Unions under Imperfect Competition: Open Regionalism Is Good. “Journal of International Economics”, vol. 41, pp. 153-177.
54. Yi, S.-S. (1997), Stable Coalition Structures with Externalities. “Games and Economic Behavior”, vol. 20, pp. 201-237.
55. Yi, S.-S. (2000), Free Trade Areas and Welfare: An Equilibrium Analysis. “Review of International Economics”, vol. 8(2), pp. 336-347.
56. Zenghelis, D. and N. Stern (2009), Principles for a Global Deal for Limiting the Risks from Climate Change. “Environmental and Resource Economics”, vol. 43, pp. 307-311.

# Appendix 1

## Non-signatories' Markets

The profit of firm  $i$  located in a signatory country in market  $l$  is given by  $\pi_{li} = q_{li}(p_{li} - c - t_i)$ . Substituting the inverse demand function in equation (4) in the text, after the appropriate changes in notation, we derive the following first order condition:

$$\frac{\partial \pi_{li}}{\partial q_{li}} = a - c - t_i - (2 - \gamma)q_{li} - \gamma Q_l = 0 \iff q_{li} = \frac{1}{2 - \gamma} [a - c - t_i - \gamma Q_l] \quad (1)$$

where  $Q_l$  is the total quantity consumed in market  $l$ . The right-hand side expression in equation (1) is the replacement function of firm  $i$  ( $q_{li} = R_i(Q_l)$ ) which is strictly downward sloping, except for  $\gamma=0$  in which it is a horizontal line. It is evident that a necessary condition for positive quantities is  $a > c$ .

The profit of firm  $j$  in a non-signatory country in market  $l$  is given by  $\pi_{lj} = q_{lj}(p_{lj} - c - t_j)$  which leads to the following first order condition:

$$\frac{\partial \pi_{lj}}{\partial q_{lj}} = a - c - t_j - (2 - \gamma)q_{lj} - \gamma Q_l = 0 \iff q_{lj} = \frac{1}{2 - \gamma} [a - c - t_j - \gamma Q_l] \quad (2)$$

where the right-hand side expression in equation (2) is the replacement function of firm  $j$  ( $q_{lj} = R_j(Q_l)$ ). By summing the  $m$  first order conditions in (1) and the  $n - m$  first order conditions in (2), we derive the aggregate replacement function  $\sum_{i \in N} R_i(Q_l)$ :

$$\sum_{i \in N} R_i(Q_l) := Q_l = \frac{1}{2 - \gamma} [n(a - c) - mt_i - (n - m)t_j - n\gamma Q_l] . \quad (3)$$

The aggregate replacement function is downward sloping over the entire domain and hence the equilibrium is unique. Solving (3) for  $Q_l$ , gives:

$$Q_l^* = \frac{n(a - c) - m(t_i^* - t_j^*) - nt_j^*}{(n - 1)\gamma + 2} . \quad (4)$$

By substituting  $Q_l^*$  in (1) and (2), we derive  $q_{li}^*$  in (11) and  $q_{lj}^*$  in (12) in the text.

### **Signatories' Markets:**

The procedure is similar as explained above. The profit of firm  $i$  in market  $k$  is given by  $\pi_{ki} = q_{ki}(p_{ki} - c - t_i)$  which leads to the following first order condition:

$$\frac{\partial \pi_{ki}}{\partial q_{ki}} = a - c - t_i - (2 - \gamma)q_{ki} - \gamma Q_k = 0 \iff q_{ki} = \frac{1}{2 - \gamma} [a - c - t_i - \gamma Q_k]. \quad (5)$$

The profit of firm  $j$  in market  $k$  is given by  $\pi_{kj} = q_{kj}(p_{kj} - c - t_j - \phi(t_i - t_j))$ , which gives:

$$\frac{\partial \pi_{kj}}{\partial q_{kj}} = a - c - t_j - \phi(t_i - t_j) - (2 - \gamma)q_{kj} - \gamma Q_k = 0 \quad (6)$$

$$\iff q_{kj} = \frac{1}{2 - \gamma} [a - c - t_j - \phi(t_i - t_j) - \gamma Q_k]$$

Summing up the first order conditions in (5) and (6), gives  $Q_k^*$  in (16) in the text, which upon substituting in (5) and (6) above, gives  $q_{ki}^*$  in (14) and  $q_{kj}^*$  in (15) in the text. in the text, (17) is derived by summing (11) over all  $n - m$  non-signatory markets and (14) over all  $m$  signatory markets. Similarly, (18) is derived by summing (12) over all  $n - m$  signatory markets and (15) over all  $m$  non-signatory markets.



# Appendix 2

No-BCA-Regime

General Results: For all coalition sizes, non-signatories' welfare is (strictly) larger than signatories' welfare. Every enlargement of coalitions from  $m$  to  $m+1$  generates strictly positive externalities for non-signatories (full strict positive externality) and is associated with full cohesiveness. This can be shown analytically. For  $\gamma = \{0,1\}$ , see Finus/Al Khourdajie (2018); results for  $\gamma = 0.5$  are available upon request.

**Table 1: Superadditivity\***

	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
$\gamma = 1$	$a_1(\delta, \gamma)$	-	-	-	-	-	-	-	-	+
	$a_2(\delta, \gamma)$	-	-	-	-	-	-	-	-	+
	$a_3(\delta, \gamma)$	-	-	-	-	-	-	-	-	+
	$a_4(\delta, \gamma)$	-	-	-	-	-	-	-	-	+
	$a_5(\delta, \gamma)$	-	-	-	-	-	-	-	-	+
	$a_6(\delta, \gamma)$	-	-	-	-	-	-	-	-	+
$\gamma = 0.5$	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
$\gamma = 0$	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_6(\delta, \gamma)$	+	+	+	+	+	+	+	+	+

\*    +: holds strictly. -: fails. Can be proved analytically. For  $\gamma = \{0,1\}$ , see Finus/Al Khourdajie (2018); results for  $\gamma = 0.5$  are available upon request.

**Table 2: Internal Spillovers of Enlargement of Coalitions\***

	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
$\gamma = 1$	$a_1(\delta, \gamma)$	-	-	-	-	+	+	+	+	+
	$a_2(\delta, \gamma)$	-	-	-	-	+	+	+	+	+
	$a_3(\delta, \gamma)$	-	-	-	-	+	+	+	+	+
	$a_4(\delta, \gamma)$	-	-	-	-	+	+	+	+	+
	$a_5(\delta, \gamma)$	-	-	-	-	+	+	+	+	+
	$a_6(\delta, \gamma)$	-	-	-	-	+	+	+	+	+
$\gamma = 0.5$	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
$\gamma = 0$	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_6(\delta, \gamma)$	+	+	+	+	+	+	+	+	+

\*      +: (strictly) positive. -: (strictly) negative internal spillovers.

**Table 3: Stability Function ( $\Phi(\mathbf{m}) := W_{i \in S}(\mathbf{m}) - W_{j \in S}(\mathbf{m} - \mathbf{1})$ ) \***

	Parameter a	Coalition Size								
		2	3	4	5	6	7	8	9	10
$\gamma = 1$	$a_1(\delta, \gamma)$	-	-	-	-	-	-	-	-	-
	$a_2(\delta, \gamma)$	-	-	-	-	-	-	-	-	-
	$a_3(\delta, \gamma)$	-	-	-	-	-	-	-	-	-
	$a_4(\delta, \gamma)$	-	-	-	-	-	-	-	-	-
	$a_5(\delta, \gamma)$	-	-	-	-	-	-	-	-	-
	$a_6(\delta, \gamma)$	-	-	-	-	-	-	-	-	-
$\gamma = 0.5$	Parameter a	Coalition Size								
		2	3	4	5	6	7	8	9	10
	$a_1(\delta, \gamma)$	+	-	-	-	-	-	-	-	-
	$a_2(\delta, \gamma)$	+	-	-	-	-	-	-	-	-
	$a_3(\delta, \gamma)$	+	-	-	-	-	-	-	-	-
	$a_4(\delta, \gamma)$	+	-	-	-	-	-	-	-	-
	$a_5(\delta, \gamma)$	+	-	-	-	-	-	-	-	-
	$a_6(\delta, \gamma)$	+	-	-	-	-	-	-	-	-
$\gamma = 0$	Parameter a	Coalition Size								
		2	3	4	5	6	7	8	9	10
	$a_1(\delta, \gamma)$	+	+	-	-	-	-	-	-	-
	$a_2(\delta, \gamma)$	+	+	-	-	-	-	-	-	-
	$a_3(\delta, \gamma)$	+	+	-	-	-	-	-	-	-
	$a_4(\delta, \gamma)$	+	+	-	-	-	-	-	-	-
	$a_5(\delta, \gamma)$	+	+	-	-	-	-	-	-	-
	$a_6(\delta, \gamma)$	+	+	-	-	-	-	-	-	-

\* +: strictly positive. -: negative. Can be proved analytically. For  $\gamma = \{0,1\}$ , see Finus/Al Khourdajic (2018); results for  $\gamma = 0.5$  are available upon request.

**Table 4: Global Welfare of Stable Coalitions\***

	Parameter a	m*	$\delta = 10$	$\delta = 50$	$\delta = 100$
			CGI(m*) %	CGI(m*) %	CGI(m*) %
$\gamma = 1$	$a_1(\delta, \gamma)$	1	0	0	0
	$a_2(\delta, \gamma)$	1	0	0	0
	$a_3(\delta, \gamma)$	1	0	0	0
	$a_4(\delta, \gamma)$	1	0	0	0
	$a_5(\delta, \gamma)$	1	0	0	0
	$a_6(\delta, \gamma)$	1	0	0	0
$\gamma = 0.5$	Parameter a	m*	$\delta = 10$	$\delta = 50$	$\delta = 100$
			CGI(m*) %	CGI(m*) %	CGI(m*) %
	$a_1(\delta, \gamma)$	2	1	1	1
	$a_2(\delta, \gamma)$	2	1	1	1
	$a_3(\delta, \gamma)$	2	1	1	1
	$a_4(\delta, \gamma)$	2	1	1	1
	$a_5(\delta, \gamma)$	2	1	1	1
	$a_6(\delta, \gamma)$	2	1	1	1
$\gamma = 0$	Parameter a	m*	$\delta = 10$	$\delta = 50$	$\delta = 100$
			CGI(m*) %	CGI(m*) %	CGI(m*) %
	$a_1(\delta, \gamma)$	3	7.3	7.3	7.3
	$a_2(\delta, \gamma)$	3	7.3	7.3	7.3
	$a_3(\delta, \gamma)$	3	7.3	7.3	7.3
	$a_4(\delta, \gamma)$	3	7.3	7.3	7.3
	$a_5(\delta, \gamma)$	3	7.3	7.3	7.3
	$a_6(\delta, \gamma)$	3	7.3	7.3	7.3

\* m\* is the same for open and exclusive membership and for a simultaneous and sequential coalition formation process.

# BCA-Regime



General Results: Every enlargement of coalitions from  $m$  to  $m+1$  is associated with superadditivity. Signatories' welfare is always higher with BCAs than without BCAs, which is not always the case for non-signatories; see Table 11.

**Table 5: Welfare of Signatories minus Welfare of Non-Signatories\***

	Parameter a	Coalition Size							
		2	3	4	5	6	7	8	9
$\gamma = 1$	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_6(\delta, \gamma)$	+	+	+	+	+	+	+	+
$\gamma = 0.5$	Parameter a	Coalition Size							
		2	3	4	5	6	7	8	9
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+
$\gamma = 0$	Parameter a	Coalition Size							
		2	3	4	5	6	7	8	9
	$a_1(\delta, \gamma)$	+	-	-	-	-	-	-	-
	$a_2(\delta, \gamma)$	+	+	+	-	-	-	-	-
	$a_3(\delta, \gamma)$	+	+	+	-	-	-	-	-
	$a_4(\delta, \gamma)$	+	+	+	-	-	-	-	-
	$a_5(\delta, \gamma)$	+	+	+	-	-	-	-	-
	$a_6(\delta, \gamma)$	+	+	+	+	-	-	-	-

\* Difference +: (strictly) positive; -: negative.

**Table 6: Externality to Non-signatories of Enlargement of Coalitions\***

	Parameter a	Enlargement of Coalition							
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9
$\gamma = 1$	$a_1(\delta, \gamma)$	+	+	+	+	+	+	-	-
	$a_2(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_3(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_4(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_5(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_6(\delta, \gamma)$	-	-	-	-	-	-	-	-
$\gamma = 0.5$	Parameter a	Enlargement of Coalition							
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	-
	$a_2(\delta, \gamma)$	+	-	-	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	-	-	-	-	-	-	-
	$a_4(\delta, \gamma)$	-	-	-	-	-	-	-	+
	$a_5(\delta, \gamma)$	-	-	-	-	-	-	-	+
	$a_6(\delta, \gamma)$	-	-	-	-	-	-	-	+
$\gamma = 0$	Parameter a	Enlargement of Coalition							
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+
	$a_6(\delta, \gamma)$	+	+	+	+	+	+	+	+

\* +: (strictly) positive; -: negative externality.

**Table 7: Internal Spillovers of Enlargement of Coalitions\***

	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
$\gamma = 1$	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	-
	$a_2(\delta, \gamma)$	+	+	+	+	+	-	-	-	-
	$a_3(\delta, \gamma)$	+	+	+	+	+	-	-	-	-
	$a_4(\delta, \gamma)$	+	+	+	+	-	-	-	-	-
	$a_5(\delta, \gamma)$	+	+	+	+	-	-	-	-	-
	$a_6(\delta, \gamma)$	+	+	+	+	-	-	-	-	-
$\gamma = 0.5$	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	-
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	-	-
	$a_3(\delta, \gamma)$	+	+	+	+	+	-	-	-	-
	$a_4(\delta, \gamma)$	+	+	+	+	+	-	-	-	-
	$a_5(\delta, \gamma)$	+	+	+	+	-	-	-	-	-
	$a_6(\delta, \gamma)$	+	+	+	+	-	-	-	-	-
$\gamma = 0$	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_6(\delta, \gamma)$	+	+	+	+	+	+	+	+	+

\*    +: (strictly) positive. -: negative internal spillovers.

**Table 8: Cohesiveness\***

	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
$\gamma = 1$	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	-	-	-	-	+
	$a_3(\delta, \gamma)$	+	+	+	-	-	-	-	+	+
	$a_4(\delta, \gamma)$	+	+	-	-	-	-	-	+	+
	$a_5(\delta, \gamma)$	+	-	-	-	-	-	-	+	+
	$a_6(\delta, \gamma)$	-	-	-	-	-	-	-	+	+
$\gamma = 0.5$	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	- <sup>1</sup>	-	+	+	+
	$a_4(\delta, \gamma)$	+	+	-	-	-	-	+	+	+
	$a_5(\delta, \gamma)$	-	-	-	-	-	-	+	+	+
$\gamma = 0$	Parameter a	Enlargement of Coalition								
		1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_6(\delta, \gamma)$	+	+	+	+	+	+	+	+	+

\* Cohesiveness holds (strictly): +; is violated:-. <sup>1</sup> holds for  $\delta = 50$  and  $\delta = 100$  but is positive for  $\delta = 10$ .

**Table 9: Stability Function  $\Phi(\mathbf{m}) := W_{i \in S}(\mathbf{m}) - W_{j \notin S}(\mathbf{m} - 1)$ \***

	Parameter a	Coalition Size								
		2	3	4	5	6	7	8	9	10
$\gamma = 1$	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_6(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
$\gamma = 0.5$	Parameter a	Coalition Size								
		2	3	4	5	6	7	8	9	10
	$a_1(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
	$a_6(\delta, \gamma)$	+	+	+	+	+	+	+	+	+
$\gamma = 0$	Parameter a	Coalition Size								
		2	3	4	5	6	7	8	9	10
	$a_1(\delta, \gamma)$	+	+	+	+	+	-	-	-	-
	$a_2(\delta, \gamma)$	+	+	+	+	+	+	+	-	-
	$a_3(\delta, \gamma)$	+	+	+	+	+	+	+	-	-
	$a_4(\delta, \gamma)$	+	+	+	+	+	+	+	+	-
	$a_5(\delta, \gamma)$	+	+	+	+	+	+	+	+	-
	$a_6(\delta, \gamma)$	+	+	+	+	+	+	+	+	-

\*   +: (strictly) positive. -: negative.

**Table 10: Global Welfare of Stable Coalitions\***

	Parameter a	m*	$\delta = 10$	$\delta = 50$	$\delta = 100$
			CGI(m*) %	CGI(m*) %	CGI(m*) %
$\gamma = 1$					
	$a_1(\delta, \gamma)$	10/9-10/9	100/98.7	100/99	100/98.7
	$a_2(\delta, \gamma)$	10/6-10/6	100/93.5	100/93.4	100/93.5
	$a_3(\delta, \gamma)$	10/6-10/6	100/44.1	100/43	100/44.1
	$a_4(\delta, \gamma)$	10/5-10/5	100/<0	100/<0	100/<0
	$a_5(\delta, \gamma)$	10/5-10/5	100/<0	100/<0	100/<0
	$a_6(\delta, \gamma)$	10/5-10/5	100/<0	100/<0	100/<0
$\gamma = 0.5$	Parameter a		$\delta = 10$	$\delta = 50$	$\delta = 100$
		m*	CGI(m*) %	CGI(m*) %	CGI(m*) %
	$a_1(\delta, \gamma)$	10/9-10/9	100/97.3	100/97.5	100/97.6
	$a_2(\delta, \gamma)$	10/8-10/8	100/95.5	100/95.3	100/95.2
	$a_3(\delta, \gamma)$	10/6-10/6	100/78.7	100/74.2	100/74.1
	$a_4(\delta, \gamma)$	10/6-10/6	100/<0	100/<0	100/<0
	$a_5(\delta, \gamma)$	10/5-10/5	100/<0	100/<0	100/<0
	$a_6(\delta, \gamma)$	10/5-10/5	100/<0	100/<0	100/<0
$\gamma = 0$	Parameter a		$\delta = 10$	$\delta = 50$	$\delta = 100$
		m*	CGI(m*) %	CGI(m*) %	CGI(m*) %
	$a_1(\delta, \gamma)$	6/6/6	87.9/87.9	87.3/87.3	87.2/87.2
	$a_2(\delta, \gamma)$	8/8/8	97.7/97.7	97.4/97.4	97.4/97.4
	$a_3(\delta, \gamma)$	8/8/8	98.0/98.0	97.9/97.7	97.7/97.7
	$a_4(\delta, \gamma)$	9/9/9	99.8/99.8	99.7/99.7	99.7/99.7
	$a_5(\delta, \gamma)$	9/9/9	99.9/99.9	99.8/99.8	99.8/99.8
	$a_6(\delta, \gamma)$	9/9/9	99.9/99.9	99.9/99.9	99.9/99.9

\* m\*: size of stable coalition(s) for a simultaneous and sequential coalition formation process under open membership/simultaneous coalition formation process under exclusive membership/sequential coalition formation process under exclusive membership. CGI: closing the gap index:

$$\sum_{i \in N} W_i(m^*) - \sum_{i \in N} W_i(m=1) / \sum_{i \in N} W_i(m=n) - \sum_{i \in N} W_i(m=1), \text{ first entry largest and second entry CGI for the smallest stable coalition } m^*.$$

**Table 11: Welfare of Non-Signatories with BCAs minus No-BCAs\***

	Parameter a	Coalition Size							
		2	3	4	5	6	7	8	9
$\gamma = 1$	$a_1(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_2(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_3(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_4(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_5(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_6(\delta, \gamma)$	-	-	-	-	-	-	-	-
$\gamma = 0.5$	Parameter a	Coalition Size							
		2	3	4	5	6	7	8	9
	$a_1(\delta, \gamma)$	+	+	+	+	+	- <sup>1</sup>	-	-
	$a_2(\delta, \gamma)$	+	+	+	-	-	-	-	-
	$a_3(\delta, \gamma)$	+	- <sup>1</sup>	-	-	-	-	-	-
	$a_4(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_5(\delta, \gamma)$	-	-	-	-	-	-	-	-
	$a_6(\delta, \gamma)$	-	-	-	-	-	-	-	-
$\gamma = 0$	Parameter a	Coalition Size							
		2	3	4	5	6	7	8	9
	$a_1(\delta, \gamma)$	+	+	+	+	+	-	-	-
	$a_2(\delta, \gamma)$	+	+	+	+	+	-	-	-
	$a_3(\delta, \gamma)$	+	+	+	+	+	-	-	-
	$a_4(\delta, \gamma)$	+	+	+	+	+	-	-	-
	$a_5(\delta, \gamma)$	+	+	+	+	+	-	-	-
	$a_6(\delta, \gamma)$	+	+	+	+	+	-	-	-

\* Difference: (strictly) positive (+) and negative (-); <sup>1</sup> holds for  $\delta = 50$  and  $\delta = 100$  but is positive for  $\delta = 10$ .



# Tables

**Table 1: Stable Coalitions and Global Welfare under Open Membership**

	No-BCA-regime		BCA-regime	
$\gamma$	$m^*$	$CGI(m^*)$	$m^*$	$CGI(m^*)$
1	1	0	10	100
0.5	2	1	10	100
0	3	7.3	6-9	87.9-99.9

**Table 2: Stable Coalitions and Global Welfare under Exclusive Membership**

	Parameter a	No-BCA-regime		BCA-regime			
$\gamma$		$m^*$	$CGI(m^*)$ %	$m^* = m_L^* - m_H^*$ simultaneous	$m^* = m_L^*$ sequential	$CGI(m_L^*)$	$CGI(m_H^*)$
1	$a_1(\delta, \gamma)$ to $a_3(\delta, \gamma)$	1	0	9-10/6-10/6-10	9/6/6	98.7/93.5/44.1	100/100/100
	$a_4(\delta, \gamma)$ to $a_6(\delta, \gamma)$	1	0	5-10/5-10/5-10	5/5/5	<0/<0/<0	100/100/100
0.5	$a_1(\delta, \gamma)$ to $a_3(\delta, \gamma)$	2	1	9-10/8-10/6-10	9/8/6	97.3/95.5/78.9	100/100/100
	$a_4(\delta, \gamma)$ to $a_6(\delta, \gamma)$	2	1	6-10/5-10/5-10	6/5/5	<0/<0/<0	100/100/100
0	$a_1(\delta, \gamma)$ to $a_3(\delta, \gamma)$	3	7.3	6/8/8	6/8/8	87.9/97.7/98.0	87.9/97.7/98.0
	$a_4(\delta, \gamma)$ to $a_6(\delta, \gamma)$	3	7.3	9/9/9	9/9/9	99.8/99.9/99.9	99.8/99.9/99.9

No-BCA-regime:  $m^*$ : size of stable coalition for a simultaneous and sequential coalition formation process; BCA-regime: entries for different values of  $a_k(\delta, \gamma)$ ,  $k=1, 2, \dots, 6$  as outlined in Section 4;  $m^*$  simultaneous (sequential): size of stable coalition(s) for a simultaneous (sequential) coalition formation process;  $CGI(m_L^*)$  ( $CGI(m_H^*)$ ) CGI for the smallest (largest) stable coalition  $m_L^*$  ( $m_H^*$ ) in the simultaneous coalition formation process, with  $m^* = m_L^* - m_H^*$  ( $m^* = m_L^*$ ) for a simultaneous (sequential) formation process. CGI for the BCA-regime assumes the particular value  $\delta=10$  with similar results for  $\delta=50$  and  $\delta=100$  (see Table 10 in Appendix 2).

Figure 1a: Stability Function and BCAs

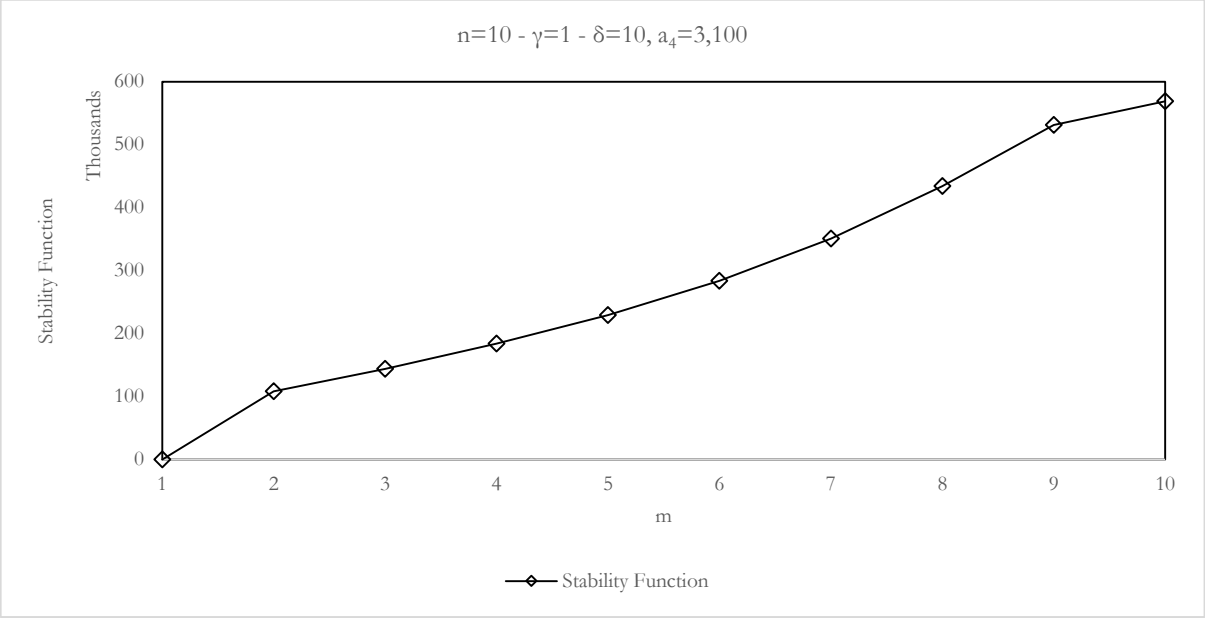


Figure 1b: Welfare of Signatories and BCAs

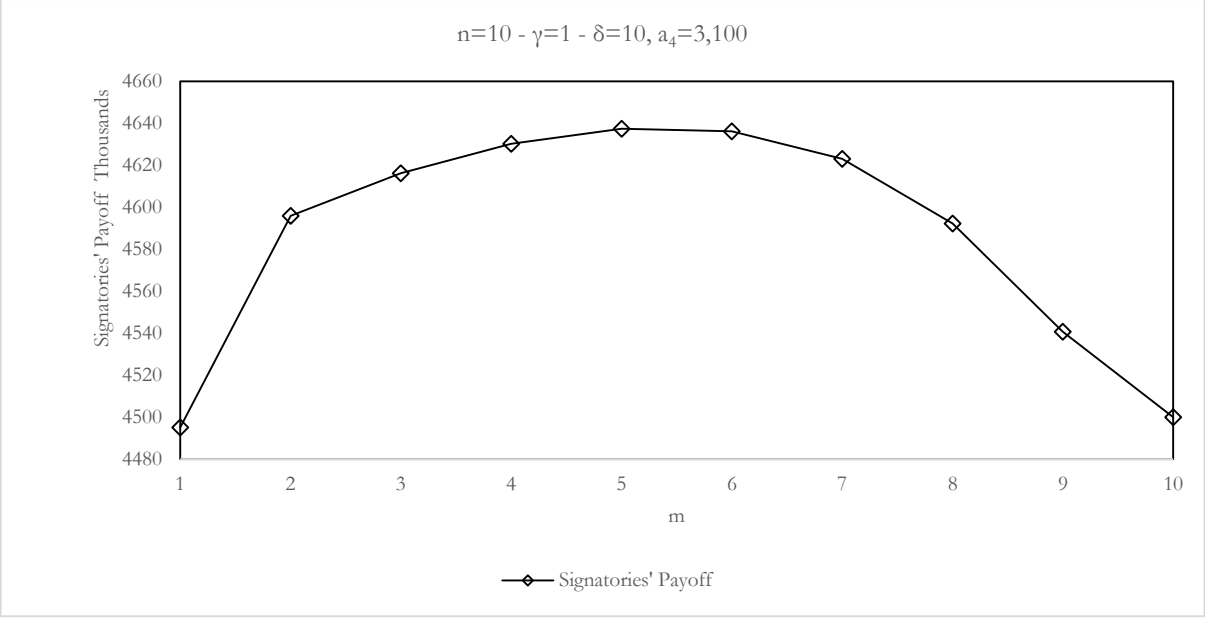


Figure 1c: Global Welfare and BCAs

